DOCUMENT RESUME

ED 447 098 SP 039 574

AUTHOR Ball, Deborah Loewenberg; Camburn, Eric; Correnti, Richard; Phelps, Geoffrey; Wallace, Raven

TITLE New Tools for Research on Instruction and Instructional Policy: A Web-Based Teacher Log. A CTP Working Paper.
INSTITUTION Center for the Study of Teaching and Policy, Seattle, WA.
SPONS AGENCY National Inst. on Educational Governance, Finance, Policymaking, and Management (ED/OERI), Washington, DC.
REPORT NO W-99-2
PUB DATE 1999-12-00
NOTE 61p.
CONTRACT R308B70003
PUB TYPE Reports - Research (143)
EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Computer Uses in Education; Data Collection; Educational Research; *Educational Technology; Elementary Education; Elementary School Mathematics; Elementary School Teachers; Mathematics Instruction; Reading Instruction; Research Methodology; *World Wide Web

ABSTRACT

This paper discusses the initial development and testing of a Web-based instrument for collecting daily data on instruction. This teacher log was developed for use in the Study of Instructional Improvement, a longitudinal study on school improvement in high poverty areas. The researchers wanted to further develop the potential of teacher logs by experimenting with Web-based technology and computer branching. The 1998 teacher log study tested the feasibility of using a Web based instrument to collect data on instruction. Seven elementary teachers used the log to report on 29 mathematics and reading lessons. Researchers observed classroom instruction, completed log reports, and wrote detailed narrative descriptions for 24 lessons. All teachers were able to use the log to report on instruction. Data were successfully entered through a Web interface into a database and subsequently used to conduct provisional analyses of both instruction and instrument reliability and validity. The researchers conclude that a Web-based instrument can be used successfully to capture important information about teaching, some of which has not been available before (lesson duration, student attendance for each lesson, and instructional groupings or formats for each lesson). An appendix offers teacher log pilot help pages. (SM)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 447 098

New Tools for Research on Instruction and Instructional Policy: A Web-based Teacher Log

A CTP Working Paper

by
Deborah Loewenberg Ball
Eric Camburn
Richard Correnti
Geoffrey Phelps
Raven Wallace

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

December 1999
(Document W-99-2)

ctp
Center for the Study of Teaching and Policy
UNIVERSITY OF WASHINGTON

SP039574

BEST COPY AVAILABLE

Center for the Study of Teaching and Policy

A National Research Consortium

UNIVERSITY OF WASHINGTON (lead institution)

STANFORD UNIVERSITY

TEACHERS COLLEGE/COLUMBIA UNIVERSITY

UNIVERSITY OF MICHIGAN

UNIVERSITY OF PENNSYLVANIA

Other active participants in CTP's research and dissemination program include researchers affiliated with Indiana University, Pennsylvania State University, the University of California at Santa Barbara, the University of Georgia, the University of North Carolina, and the University of Pittsburgh.

The Center's program of research is carried out in collaboration with various other research organizations, among them other OERI-funded research centers, including the Consortium for Policy Research in Education (CPRE), the Center for Research on Education, Diversity, and Excellence (CREDE), and the Center for English and Language Achievement (CELA). The Center is affiliated with a variety of professional and advocacy organizations which represent teachers, teacher educators, state and local policymakers, disciplinary groups, and educational reform interests.

The work reported herein was supported under the Educational Research and Development Centers Program, PR/Award Number R308B70003, as administered by the National Institute on Educational Governance, Finance, Policymaking and Management, Office of Educational Research and Improvement (OERI), U.S. Department of Education. However, the contents do not necessarily represent the positions or policies of the National Institute, OERI, or the U.S. Department of Education, or the endorsement of the federal government.

The authors gratefully acknowledge Andy Porter's helpful comments on an earlier draft of this report, as well as his previous work and ongoing counsel on the overall endeavor of designing the instructional log.

Table of Contents

Abstract 4

Introduction 5

Instrument Design 11

Methods 14

Results 15

Discussion 30

Epilogue 33

Endnotes 34

Appendix 36

ABSTRACT

This working paper reports on the initial development and pilot testing of a Web-based instrument designed to collect daily data on instruction. This instrument, referred to as the teacher log, was being developed for use in the Study of Instructional Improvement, a large scale, longitudinal study focusing on school improvement in high poverty schools. Although the instructional log we are ultimately using in this research is not a Web-based tool, and its features are both similar to and different from the one under discussion in this report, we think that there are elements of our early instrument design work that may be of use to others interested in the development of such tools to track instruction.

In recent years, a number of studies seeking to generate more detailed data on instruction have pioneered the use of teacher logs. Such logs are generally self-administered instruments which ask teachers to report on topics covered, pedagogy, and more. Teachers fill out such logs on a daily basis and data from these daily reports is then aggregated to create portraits of content emphasis and pedagogy over time.

In this pilot work, we set out to further develop the potential of teacher logs by experimenting with Web-based technology and computer branching. These features open up possibilities for capturing a much wider range of data and for linking this data to the work of individual students in classrooms. Furthermore, where paper and pencil logs must be distributed to teachers and data from each log entered into a data base, Web technology allows data to be entered directly as teachers complete the log, potentially reducing the burden on both teachers and researchers.

The teacher log pilot study was conducted in the spring of 1998 to test the feasibility of using a Web based instrument to collect data on instruction. Seven teachers in two schools used the teacher log to report on 29 lessons in mathematics and reading. In addition, project researchers observed classroom instruction, filled out log reports, and wrote detailed narrative descriptions for 24 of these lessons.

Results from this pilot were encouraging. All teachers were able to use the teacher log to report on instruction. Data was successfully entered through a Web interface into a data base and subsequently used to conduct provisional analyses of both instruction and the reliability and validity of the instrument.

INTRODUCTION

Research on instruction and on programs or policies aimed at improving instruction confronts several enduring challenges, three of which are especially important in studying instructional improvement interventions. First, such studies must capture instruction efficiently across large numbers of classrooms and over a sufficient period of time to document the unfolding of teaching and learning. Second, these studies need to portray instruction in terms that can be linked to evidence of student achievement. Third, the process of gathering such data must not intrude excessively on teachers' work in and out of the classroom. Teachers' collaboration in this process is essential, and, to the extent that teachers are providers of the data, their interest and understanding of the research instruments are also important.

Among the various ways of gathering data on classroom instructionæincluding third-party observation, interviews of teachers and students, and analysis of student work productsæteacher logs hold special promise. Filled out by teachers on a daily or periodic basis, teacher logs offer a way to capture many facets of instruction in an efficient and grounded manner, through brief, prompted surveys. Pioneering work on teacher logs has begun to demonstrate both the potential and limitations of this research tool. While there are shortcomings of self-reported data, work with teacher logs to date suggests that, for many aspects of instruction, this tool can help researchers construct a useful and informative record of instruction.

This working paper reports on the first stage of development and initial pilot testing of a teacher log instrument we have been developing for use in a large-scale national study of instructional improvement. The instrument represents an attempt to substantially expand the capacity of teachers' logs through the creation of interactive instrumentation, accessed by teachers over the Internet. The pilot study offered a first test of instrument design assumptions and feasibility while helping to uncover issues confronting further design work and pilot testing.

To situate discussion of the log, we begin with a synopsis of the Study of Instructional Improvement in which this instrument will be used. We continue with a description of the pilot we conducted during May and June, 1998, trying out a first version of an instructional log. The report includes the main questions we pursued in the pilot, a description of the draft log instrument itself, and details about method and setting of the pilot. We continue with a discussion of what we learned from the pilot, including basic analyses of the data, as well as some of the conceptual, practical, and technical results.

Study Context

The Study of Instructional Improvement at the University of Michigan is focusing on school improvement in high poverty elementary schools. Our inquiry is motivated by two developments: (a) an increase in the number of promising whole-school instructional interventions, and (b) growing interest in tying public funds to evidence of improved student performance.[1] The study will examine the effects of a few of these school improvement interventions on principals, teachers and students. Specifically, the study will document those features of interventions which seem likely to influence instruction and learning. In addition, we hope to better understand factors which are likely to mediate the enactment of any intervention. These include organizational, professional, and environmental attributes—everything from local community characteristics to state and federal policy—which may

influence instruction and learning. Our theoretical frame implies that in order to understand efforts to change instruction, we must focus on three distinct types of interactions. Among these are interactions within schools and their own instructional organization, interactions between people in schools and change agents seeking instructional improvement in those schools, and interactions in the larger environment in which both schools and change agents operate.

The study's key questions include:

- What elements in the design of school improvement interventions are most likely to contribute to improved instruction and student performance?
- What features of state and local policy environments are most likely to support or impede effective intervention?
- What features of intervention implementation—including community involvement, instructional coordination, new instructional roles, leadership, or teachers' opportunities to learn—bear most directly on instruction and student learning?
- What are the costs—in human and social resources, political commitment, organization, and money—of effective intervention implementation and improved student performance?

The study is an ambitious undertaking in data collection. It is designed to be a six-year longitudinal study in approximately 125 schools. About 100 of the schools will be involved in whole-school change interventions. The other 25 schools will be selected because they are not involved in an intervention.[2] In addition, we will conduct intensive case studies on issues key to the study in a number of schools within each intervention. The case studies will permit us to probe issues such as intervention dynamics and teacher learning in greater detail than we could in the larger sample of schools. Also, because the same instruments will be used with both sets of schools, the case studies will enable us to cross-validate the data from the larger sample.

In addition to questions about the nature and effects of whole school interventions, the study will scrutinize underlying educational processes. For example, our data will enable us to associate evidence of students' exposure to educational resources with evidence of their academic and social performance during the six year time period. This will allow for more precise estimates of the following factors:

- The effects of particular educational resourcesæsuch as, time spent on academic tasks, teachers' pedagogical and content knowledge and instructional coordination æ individually and in relation to one another.
- The cumulative effects of such resources.
- The costs of such resources.

In order to collect data on students, teachers and schools, the study will use a variety of instruments and approaches for data collection. Among these methods are, surveys, interviews, observations, and collection of work artifacts (i.e., curriculum materials and student work). Instructional data is crucial to answering questions in the study. This data will be collected directly from teachers in a variety of ways. One central tool, which we refer to as the *teacher log*, will elicit from teachers a daily report of their mathematics and reading instruction.

Using the teacher log, hereafter referred to as *the log*, we aim to get a detailed picture of instruction for students as they pass through successive grades in each of the schools under study. We aim to create a tool that will take teachers no longer than 15 minutes to complete for mathematics and reading instruction combined.

We see three main purposes for collecting these data on instruction:

- One aim is to develop a more complete picture of what teachers do in teaching math and reading than currently exists. There is limited empirical evidence about the modal patterns of instruction in mathematics and reading. Little is known about the (variability in) content that teachers teach, how they approach it, how much time they allocate to which topics, the materials they use, how they group students for instruction, etc.
- A second purpose for collecting information about classroom instruction is to analyze the impact of interventions on teachers' practices. Some intervenors design explicit opportunities for teacher learning while others aim to influence students' school experiences through less direct means. We want to investigate how the different interventions influence classroom instruction. In this sense, instruction is a dependent variable in analyses of interventions, as we seek to understand how teachers' and schools' opportunities to learn are related to changes in teachers' practice.
- A third purpose is to associate more detailed evidence of instruction with evidence of student performance. In this sense, instruction is seen as an independent variable for analyses of student learning. For example, we will be able to collect longitudinal information about the time students spend on different kinds of academic tasks in reading and mathematics, the content to which they have access, and the instructional formats and approaches that characterize their learning opportunities. We can then relate that evidence to longitudinal data on growth in students' academic performance, a critical link in the effort to understand the impact of school improvement efforts.

Specific Questions for the Log Pilot

The questions that guided the pilot work for the teacher log fall into three general categories: (a) questions about conceptualization and measurementæthat is, about the articulation and framing of those aspects of practice that we aim to capture and the validity and reliability of the data we gather; (b) practical questions about how to design a log that has the potential to gather detailed instructional data and that teachers will be willing to fill out on a daily basis; and (c) technical questions about the feasibility of developing a Web-based instrument. The pilot helped further our thinking in relation to several of these original questions while leading to a host of new questions. In this section, we give a brief overview of the questions that shaped the pilot study.

Questions about Conceptualization and Measurement

Previous efforts to log teaching practice—in particular, the work of Andrew Porter and his colleagues, as well as the work on curriculum indicators by Lee Burstein and Lorraine McDonnell—have attended principally to content coverage, and, to a lesser extent, how content is represented to students. Subsequent work has extended the use of teacher logs to some aspects of pedagogy, characterized at the classroom level.[3] However, since we seek to document the content to which each individual student is exposed, obtaining lists of topics covered on a given day or broad characterizations of pedagogical approach is insufficient for our purposes. We need to be able to link topic coverage with individual student activity on a lesson by lesson basis. There are many potential sources of variation in classroom instruction. These include, for example, variation in content delivered to particular students, variation due to grouping practices, materials or representations used to deliver content to different students, and much more. While it is clear that individual students might have quite different instructional experiences within a classroom depending on how they are brought into contact with the subject matter, it is not clear how to best capture this variation. One of the primary purposes of this research is to

advance understanding of the problems of capturing the instructional experience of individual students across the many possible ways teachers organize classroom instruction.

One of the primary problems is developing valid methods for participating teachers to divide the school day into discrete blocks of time. Some of these blocks of instructional time may conventionally be conceived of as 'lessons'. However, 'lessons' don't appear to be an agreed-upon construct. What one teacher thinks of as a lesson may be very different from another teacher's definition. This is further complicated by the fact that many teachers teach mathematics and reading in small increments at several points during the day. Furthermore, in some lessons mathematics and reading are taught concurrently with some students working on mathematics while other students work on reading. Moreover, we need records of instruction that account for within-lesson shifts in topic, grouping, instructional approach, and student activity.

Another problem we face is developing valid content maps that allow teachers to accurately portray the content of instruction. Most previous attempts to gather data on content coverage through teacher logs have focused on mathematics and science. Relatively little work has been done to map the territory of reading and language arts instruction. (See endnote 3 for one example.) Whereas in mathematics we have previous work to draw from, in language arts we face the additional problem of creating useful maps of the subject area. Finally, we must also map language arts in ways that are congruent with teacher conceptions, instructional practice, curriculum materials, and assessments.[4]

In both language arts and mathematics, the problems of developing content maps and ways of reporting the details of instruction take us into areas which are highly inferential and about which there is little consensus on the most common language for describing content and on the way this content is organized in instructional settings. The language used is often value-laden and lends itself to multiple interpretations. The presentation of questions on the log—from the exact wording of questions to the form and structure assumed for lessons—may be important determinants of how teachers interpret the log, and thus, how they report on their teaching. Since so little is known about how teachers conceptualize content, we used this first log as an opportunity to learn more about how teachers define the topical content of lessons.

In this pilot, we were interested in learning about the following:

- What problems arise as a result of having teachers report on lesson-level "chunks" of instruction in reading and mathematics?
- In designing log items to which teachers would respond, what language can we use to discuss topics of instruction, instructional treatment of these topics, and the organization of students for instruction?
- Is it feasible to collect data on instruction that will allow us to characterize variation in the instructional experience of individual students in the classroom?
- Is it possible to design an instrument that is intuitive and has construct validity for teachers?
- What other issues are raised in framing log items? Are there significant aspects of instruction which we miss using the items on the pilot?

In considering questions of conceptualization and measurement, we are also concerned about the validity and reliability of the instruments and about the process we might take to validate data collected by this and subsequent versions of the log. Therefore, we used this first pilot to both gather evidence

regarding the accuracy of teachers' responses to log items and to determine the kind of data needed to conduct a proper validation of improved, future versions of the log.

For this pilot, we considered the following questions:

- What are valid ways of characterizing the topical content of instruction?
- What was the validity of the specific data items included in the pilot?

Practical Questions about Feasibility, Burden, and Incentives

A major issue for the development of the teacher log is whether teachers can and will complete the log. It is not unusual for educational researchers to assume that teachers are not only disinterested in research but are highly unlikely to engage in research activities that require great thought or great effort. One explanation provided to support this common assumption is that teachers are so busy that any additional work will be seen as intrusion. Less generous explanations suggest that teachers are disinterested in work that requires deep and careful reflection on their teaching practice. Both of these assumptions about teachers and their willingness to engage in research efforts raise serious questions about the feasibility of complex instruments such as the teacher log.

Rather than assuming that teachers would reject an instrument like the log outright, we entered into this research seeking to understand how teachers might take up this work and what might serve as powerful incentives for their involvement. There are many possibilities why teachers might have resisted or avoided involvement in past research efforts. Some teachers may have been deterred by the design of the research and the instruments being used, while others may have been dissuaded by the conduct of the researchers themselves. It is possible that many past efforts to study instruction have characterized the instructional work of teachers and students in ways that make little sense to teachers, in essence rupturing and pulling apart the features of instruction which give it coherence. One goal of this study was to determine the feasibility of an instrument that sets out to provide teachers with a way to record some of the complexities of how they work with different students, both over time and within lessons, in ways that they see as valid. We hypothesize that teachers might find an instrument that provides a concise record of their instructional practice compelling and useful, thus providing a powerful incentive for their involvement.

We also understand that the log will need to be accompanied by strong incentives, since the burden on teachers is likely to be high. Teachers may be asked to complete parts of the log during short sampled periods of the school year, while other parts of the log would be completed on a daily basis across the entire course of the school year. Given the design of this study, using log instruments to document curriculum and instruction will require teachers to complete as many as 500 log entries over the course of the research project. This brings the question of burden to the fore in considering the viability of a log instrument. For the log to be feasible it must be simple to complete and become part of a teacher's daily routine. Strong teacher incentives coupled with routines for checking log entries may be needed to encourage regular and reliable task completion.

A second question has to do with management of the log. For the full study (over six years), the log might be filled out by as many as 900 teachers on a daily basis yielding as many as 450,000 log entries for the total sample. The burden of distributing and gathering log surveys and entering and managing data also needs to be considered. One set of questions for the pilot work bears on the feasibility of asking

teachers to complete a computer instrument on a daily basis. Another practical question has to do with the possibility of reducing the burden of managing such a large number of surveys and data points over time. We considered the following questions in the pilot study:

- What kinds of burdens does completing the log impose on teachers?
- Are there intrinsic incentives for teachers to complete the log?
- What extrinsic incentives might help motivate teachers to complete the log?

Technical Questions About Web-based Log Design

Efforts to log teaching practice on a daily basis have previously relied on paper and pencil survey instruments.[5] These have been successful at documenting content coverage, but have also had some disadvantages. In order to keep the instrument to a usable length and structure, the complexity and scope of such logs has been limited. Moreover, data entry has been a significant task, constraining practical sample sizes of participating teachers. A Web-based log instrument may provide certain advantages over a paper and pencil instrument. First, computer technology allows for much more complex branching schemes which would be extremely cumbersome in paper and pencil versions. Computers also allow customization of the log for each teacher, retaining class lists, frequently used materials, student groups, and more. Because of this, we expect that we may be able to design an instrument that permits us to collect data on a wider set of elements of practice because of the visible simplicity of the tool. Second, unlike a paper-pencil survey, a Web-based interface allows for immediate data entry. When teachers fill out the log their input goes directly into the database. We suspect that a Web-based instrument will allow for greater scope in items, increased flexibility in the kinds of information that can be tapped, and the capacity to gather data from more teachers with reduced data entry needs. Our pilot work was a first step in exploring this possibility.

However, the technical issues of designing a Web interface and a database suitable for this work are entirely untested. As far as we can determine, no previous Web-based instrument exists that is designed to document particular practices or behaviors.[6] Therefore, the first trial of the log instrument addressed some basic questions:

- Is it possible to develop a working Web interface that presents teachers with relevant categories for recording their teaching practice?
- Can the database be customized for each teacher, allowing for reporting of instruction by student, student group, and materials used in a particular classroom?
- What technical problems arise in the use of the instrument?
- What are the parameters of use? For example, what kinds of classroom computer resources do they need? What sorts of technical training and ongoing support might they need?

Learning from the Log

Analysis of the data collected during the log pilot helps us begin to understand what we can learn about instruction from this instrument, and also about its practical feasibility and technical quality. In subsequent sections, we consider how well this first version of a log provided records of what we observed, about how the data might be analyzed and what those analyses would tell us. We also provide the teachers' comments on their experience using the log to create records of their teaching.

INSTRUMENT DESIGN

The teacher log on which we report here—our first effort to develop such an instrument—was a Web-based instrument, accessible through a Web browser (e.g., Netscape) on any computer connected to the Internet. Using a home or classroom computer, teachers were able to fill in the log by accessing a secure Web site. There were a number of features of the teacher log that distinguished it from a traditional paper and pencil instrument. Before using the log to record instruction, teachers were prompted to provide basic information about themselves, their students, and any standard groupings they used in instruction (e.g., reading groups). These data were captured in a database and became part of each teacher's individual record. When teachers used the log to create records of instruction, these data were accessed and teachers were presented with their own personalized instrument. For example, in one part of the log, teachers were asked to indicate if any students were absent for a given lesson. In this section of the instrument teachers were presented with a list of students in their classroom. This student list was generated from each teacher's record in the database. Teachers could then check which students were absent and the reason for the absence.

By using computer technology it was possible to design a log that takes advantage of complex branching schemes. Depending on how teachers answer key questions they were routed through different sets of subsequent questions, thus minimizing the total number of questions teachers have to consider. In other instances teachers were prompted for additional information on particular items, sometimes using drop lists and other times utilizing open text fields. Of course, traditional paper and pencil instruments can also utilize these features. However, when using a paper and pencil instrument, respondents must thumb through many pages as they follow skip patterns. An instrument like the log could easily reach the size of a small pamphlet if it were designed as a paper and pencil instrument. The computer design allowed respondents to fill in the instrument, moving seamlessly through skip patterns only answering relevant questions.

Pilot teacher log design

In this section we provide a brief overview of the teacher log design, branching schemes, and variables. Prior to using the log to report on a specific lesson, teachers entered personal and class data into the log database. These data included grade taught, any regular groupings used in mathematics, or reading, and a class list of students. These data were subsequently accessed and used to generate a personalized log reporting form for each teacher.

After teaching a lesson, teachers logged on to a Web site by entering their name and password. Teachers began by recording the total number of times (e.g., periods or lessons) that they taught mathematics and reading on a given day. Based on this initial report the computer then prompted teachers to enter data separately for each period of instruction. For example if a teacher taught two reading lessons and one math lesson, they would fill out three separate log reports, two for reading and one for mathematics.

For each lesson, teachers indicated the length of the lesson (total time in minutes). Teachers then reported on whether or not they grouped students for instruction (e.g., whole group lecture or discussion, small group, independent work). If teachers did group students, they were then asked to report whether students stayed in these groupings or moved through different groups, and whether

students in groups worked on the same topics or different topics. Teachers were also asked to report on topic treatment. If teachers taught multiple topics, they were asked to report on whether the topics were taught sequentially or concurrently and if student activity changed or stayed constant across these topics.

For each lesson, teachers also reported on teacher activity, use of student grouping if any, and work with individual students. Teachers were asked to indicate if the lesson was taught by a classroom teacher, a paraprofessional, a parent, an intervention teacher, or another teacher. If teachers indicated that an intervention teacher taught the lesson, they were further prompted to provide the teachers name and affiliation. Likewise if another teacher taught the lesson they were again prompted to provide name and affiliation (e.g., district mathematics coordinator). Teachers then reported on their activity (e.g., worked with individual students, worked with small groups, circulated around the room, lectured or led a discussion, non-instructional). If teachers used standard instructional groups they were prompted to report on these groups from their pre-entered information on grouping. Finally teachers were prompted to note any students they worked with individually, again marking students from their pre-entered class information.

Data on student attendance and activity were also collected at the lesson level. Teachers indicated any students absent for the lesson, checking students from a pre-entered class list. For each of these students, teachers provided additional information about the reason for their absence (e.g., not at school, removed from the room for behavior problems, working with a group in another teacher's classroom, working with a tutor in a small group, working with a tutor one on one, other reason). Teachers also recorded student activity for each lesson. Finally, teachers indicated the nature of the materials students worked with during the lesson. In mathematics teachers reported on the following categories: manipulatives, worksheets, flash cards, computers (teacher prompted to provide title of software), instructional games, teacher-led activity, other. In reading teachers reported on the following categories: reading trade books, working on textbooks, working on worksheets, working on flash cards, working on a computer (teacher prompted to provide title of software), playing instructional games, working on teacher led activity, other activity.

For each lesson, teachers also reported on the specific topic or topics. Teachers listed each topic in an open field using their own descriptive language. In the cases where teachers taught multiple topics in a lesson, these topics became the basis for further branching in the log instrument. For example if teachers reported teaching two topics in reading they were subsequently prompted to answer two sets of questions for each topic.

In both reading and mathematics, topic-level data was collected on the nature of the task, the purpose of the task, and student engagement for each topic. Teachers were asked to report on the nature of the task, or cognitive complexity, of each topic they reported. Nature of task categories for reading were: (a) practiced phonics and word recognition, (b) answered literal questions, (c) needed to integrate information, and (d) needed to rely on personal experience, knowledge, opinions. For each topic they reported, teachers were asked to rate (on a four point Likert scale) whether the reading categories above were most applicable to not at all applicable. Nature of task categories for mathematics were: (a) required memorizing basic facts, (b) required recall and basic relationships or definitions, (c) were computational in nature, (d) were verbal or pictorial in nature, and (e) had no clearly defined strategy

or steps. Again teachers used a similar procedure in rating the degree each of these categories was present in their lesson. For each topic, teachers also reported on the purpose of the task (e.g., introduce, further develop, review, conclude) and student engagement with the task (low, moderate, high). Finally, topic level data were also collected on the curriculum materials. Teachers recorded the title, publisher, whether the materials were selected by the teacher or student, and source of the material(s) (e.g., intervention, workshop, district curriculum materials) for each topic.

Readers interested in a more detailed portrait of the instrument should refer to the "Teacher Log Pilot Help Pages" (Appendix). The help pages were produced as a reference manual to help teachers learn how to use the log and understand the questions and categories included in the log. However, the help pages are also a useful reference for studying the pilot log instrument itself. The help pages include the actual Web screens that teachers viewed when accessing the log. The pages also include descriptions designed to walk teachers through set-up and use of the instrument and provide an explanation of how the log works from a user's perspective.

The following instructional variables were included in the pilot version of the log:

- Amount of time devoted to reading and mathematics lessons (not relative to the rest of the day).
- Content topics covered, as identified by the teachers.
- Materials used.
- Nature of instructional tasks (focus, kind of work).
- How teachers group students into ability or other groups.
- Instructional format during lessons (grouping for specific lessons).
- Teacher activity in relation to students (did the teacher work with the whole group, small groups, or individuals (and if so, which ones?).
- Other adults' roles in instruction (e.g., aides, parent volunteers).
- Student activity during lessons (what students were doing and what they were working with).
- Student absences from specific lessons, and reasons for absence.
- Teachers' general comments about each lesson (unstructured—this was an open field).

Database and web design

The software used to develop this version of the teacher log was FileMaker Pro 4.0 (a database program) with its built-in Web Companion which facilitates creation of Web pages which communicate with the database. The architecture of the FileMaker databases consists of ten tables linked relationally to store information about teachers, students, and lessons. Tables are updated as teachers enter data through the Web interface. Teachers can also access prior lesson entries by requesting reports of what they taught on previous days. The pilot log instrument was supported by a manual that gave an overview of the log, the purposes of the pilot, and a page by page description of how to enter data into the log (Appendix). This manual was provided to teachers as a paper document and also posted as Web pages.

METHODS

Seven teachers from two elementary schools participated in the pilot study, conducted during May and June 1998. The two schools were selected because their principals had previously worked with one of the project researchers and were willing to support the teachers participating in this pilot study. The principals identified teachers who they felt would be interested in working with us, and we followed up with a letter explaining the project. Teachers were offered a stipend of $100 to participate in the pilot study. The study was promoted as a joint effort with teachers, helping to pilot the preliminary instrument in their classrooms and also helping the researchers think through issues for further development. The teachers agreed to approximately eight hours of work on the project. Participation included a short training session on using the log. The teachers were then asked to use the log to report on two mathematics and two reading lessons during which they were also observed by a researcher. A researcher interviewed each teacher after all lesson data were collected. The interviews included questions about the reported lessons as well as general questions about their practice. At the end of the pilot, teachers, principals, and researchers met for a dinner and concluded with a focus group discussion.

These teachers and the contexts in which they work are probably not representative of the teachers whom we will ultimately ask to use the log. Several features distinguished them. First was their engagement with the project. These teachers were willing not only to pilot the instrument but also to engage in discussions about further development, and to do so at a particularly busy time of year, a few weeks before summer break. A second feature that distinguished this group was their attitude toward research. While many teachers are deeply skeptical about research, these teachers were interested in what we were doing and even found it useful. Finally, the context in which these teachers worked was probably atypical. These teachers had computers in their classrooms and were already fluent and comfortable with technology, including using the Internet and email. Learning to use the log was clearly facilitated by their basic comfort with computer technology.

Hence, this first pilot of the teacher log was in many ways a best case scenario. The log will ultimately have to be usable by teachers who may be much more skeptical about educational research and researchers than were these teachers, in contexts that are less supportive of computer use, with fewer resources, and with greater pressures on teachers' time and attention.

Researchers visited the classrooms of each teacher in advance of data collection to make sure they could access the Web and could view the log from their classroom computers. Each teacher in the pilot had access to a computer in his own classroom. We were able to get each teacher set up with little difficulty. Several computers needed new versions of Netscape, and, in one case, installing Netscape led to some other problems unrelated to the Web.

Before reporting on lessons, each teacher generated a class list by entering the names of all students they taught for either math or reading. Teachers who used standard groups for delivering instruction in either reading or mathematics entered group names and indicated the students in each group. Over the course of the pilot work, each teacher reported on four lessons, with the exception of one teacher who reported on eight lessons, one who reported on three lessons, and one who reported on two lessons. Thus, we received reports on a total of 29 lessons across our teacher sample. Approximately half of these lessons were in reading (n=15) while the others were in math (n=14). Within lessons, teachers

reported on topic, lesson format, the nature of student activity, individual students and groups of students receiving instruction, and instructional materials.

The data we collected from the pilot comprised:

- Log reports from seven teachers (29 lessons in all).
- Matching log reports for 24 lessons reported by a project researcher. (There were 28 reports in all. Two researchers observed simultaneously and reported separately on four of the 24 teacher-reported lessons; five teacher-reported lessons were not observed.)
- Observation reports from the 24 lessons that were logged by both the teacher and the researcher.
- Interviews with five of the seven teachers were conducted after all lessons were reported. Interviews were taped and transcribed.
- Tape and transcript of the focus group dinner, which was attended by all seven teachers. The focus group was conducted by a project researcher using an outline prepared with questions to probe the teachers' perspectives on the log, the pilot experience, ways in which the instrument might work with other teachers, and ways to improve its usefulness and incentive for other teachers.

RESULTS

We report the results of our pilot by the sets of questions noted earlier in this report. Thus, the first set of results concerns the conceptualization and measurement of instruction accompanied by observations on what we learned about instruction from the log and about the instrument's validity. The second set concerns findings regarding practical matters, and the third, the technical feasibility of a Web-based log design.

Results Regarding Conceptualization and Measurement

The log worked well as a tool for collecting basic descriptive information about instruction. Teachers had little difficulty identifying a lesson to report and provided data by lesson on low inference items such as student absence, teacher names and roles, who taught each lesson, and the names and publishers of instructional materials. In addition, although there were some discrepancies between teachers and observers, the log also captured duration of lessons (see the section on validity below for a complete discussion of differences in teacher and researcher reporting on duration). Although teachers had some difficulty in understanding the meaning of certain items, the overall branching scheme in the log presented little difficulty. Given the degree of complexity in many of the lessons teachers described, it is highly encouraging that the overall architecture of the instrument presented no major problems.

However, the information collected at the lesson level did not capture data that allowed clear linking of instruction to students. In part, this is because of the ambiguity of the definition of "lesson." It is possible that using "lesson" as the demarcation of instruction may lead to unreliable reporting. What one teacher saw as one lesson with three parts, another teacher might see as three separate lessons. Even the same teacher may divide her instruction this way at another time. In the first instance, we obtain records of topic, material, and activity information for a single block of time. In the second instance, these items are recorded with a finer metric. In this pilot, researchers were present to observe a single lesson at a time, and thus the demarcation of lesson was made clear by their presence.

The pilot work uncovered numerous areas in which our choice of item wording may have been confusing. For example, reporting on "instructional format" was problematic for some teachers due to the categories we provided for indicating student grouping (e.g., individual instruction, small group, and whole group). In more than one instance, students were given whole group instruction and then worked in pairs at their seats. It was unclear in this version of the log how such a lesson should be characterized. Does working in pairs constitute small group work? Also, if teachers marked that students were in whole class instruction and also did small group work, this version of the log did not tell us if they spent more or less time in one configuration. Teachers had similar problems describing student activity (e.g., worked in textbooks, worksheets), teacher activity (e.g., lectured, circulated around room), and items that asked them to describe how they handled the content of the lesson. We noted teacher confusion in several other instances: one teacher thought "non-instructional" referred to any activities she did which did not involve her lecturing or leading a discussion in the front of the room. In other words, she characterized walking around the room, checking on students' work, and talking with students as non-instructional. Also, one of the teachers used worksheets during several lessons, but he never checked that his students used work sheets. In part, this is a problem of carefully defining language we use in the log, but it is also a conceptual problem of determining exactly what level of detail we want and of making our conceptual distinctions clear to teachers.

Over the course of the pilot work we ran up against many "special" situations which did not fit appropriately in this version of the log. One such problem occurred as a result of students switching classrooms for particular lessons during the day. At one school, this was a routine activity since there were several multi-grade classrooms. This posed problems for data collection for certain students. This is also a potential problem for pull-out instruction. How do we keep track of students who leave a teacher's class for instruction elsewhere? In other cases we found that our narrow definition of reading instruction was not sufficient for capturing a broad range of reading-related language arts activities. We also found that certain very complicated activitiesæsuch as work centers with multiple topics and activitiesæwere far too complex to be adequately captured by the current version of the log.

Another problem came up in reporting "nature of task," by which we meant what kind of learning was entailed in the lesson. For reading, the categories were geared toward activities in which students were asked to read text. They make little sense for either writing or listening activities. In mathematics, nature-of-task items were limited to a focus on mathematical content and did not address activities directed toward other kinds of goals, such as communicating about mathematics. We may need to increase the number of categories to include a broader range of mathematics and language arts activities. However, here we run the risk of overloading teachers by asking them to consider a great number of very complicated distinctions for each topic they teach.

What the Log Revealed About Instruction

Using data from these teachers, we investigated what we could learn about their teaching. Even though the sample was small, the results provided insight into: (a) the work of particular students and student groups; (b) the nature of student activity; (c) the nature of teacher activity; and (d) the topics and materials used in instruction.

Teachers' reports on individual students and standard student groups

The pilot log linked instruction to student activity in three ways. First, teachers indicated in the log the names of any individual students with whom they worked during a lesson. Second, teachers indicated whether they worked with any previously identified student groups (e.g., reading groups). Third, teachers reported the names of students who were absent. In slightly fewer than half of the lessons, teachers identified students on their class list with whom they worked individually. Teachers were slightly more likely to report working with individual students in mathematics lessons than reading lessons. There was considerable variation in the number of individual students with whom teachers worked across the various lessons. In five of the eight math lessons in which teachers reported working with individual students, teachers identified five or fewer students to whom they gave specific attention. In contrast, in five of the seven reading lessons, teachers named seven or more students with whom they worked individually. We do not know enough about what teachers meant by "working with a student individually."

All but one of the pilot teachers used the log to report absences for individual students. Student absences were only noted in about half the total lessons.[7] Each time a teacher reported a student absent, the log prompted her to provide a reason for the absence. The log provided teachers with four reasons from which to choose:

- Receiving instruction in another teacher's class
- Receiving instruction from a tutor
- Not in school
- Other

Most of the teachers indicating an absenteeism also provided the reason for student absence. The most frequently cited reason was that students were receiving instruction in another teacher's classroom. It will likely prove instructive in the main study to probe for more information about where the students were and what instruction they received during these lessons. Are they in a pullout program? Have they simply switched classes in a school where students move between different teachers and classrooms organized into subject-matter specialties?

A few of the teachers who chose the "other" category said that the student they were reporting on had moved. This suggests that the log might be able to capture detailed information on student mobility. In our effort to redesign the log, we have discussed the possibility of providing means for teachers to document mobility among their students by prompting teachers to record any students who have left or joined the class. This would not only keep the class list up to date but would also provide accurate information on student mobility in classrooms and schools.[8]

Finally, teachers provided information at the student level by indicating when they worked with instructional subgroups. When teachers initially entered their class lists into the log, they were given the opportunity to identify standard groups of students with whom they regularly worked. The most typical example is reading groups in the primary grades, but teachers were able to identify groups formed for other reasons. For example, one teacher taught a split first/second grade classroom. She often differentiated instruction by having students in the two different grades work on different tasks. In addition to learning more about teachers' grouping practices by allowing teachers to report on

standard student groups they used, we hoped to capture the nature of learning opportunities provided to different students through these specific groups.

Only two out of the seven teachers entered standard groups (e.g., reading groups) into the log. The two teachers who did enter groups did so for reading groups but not for math groups. Both of these teachers reported on their groups at least once. In order to understand whether asking teachers to report on standard groups will be useful, there are two issues which merit further inquiry. First, it would help us to know how often teachers actually use standard groups in their day-to-day practice. Second, it would be useful if future pilot work could verify whether teachers did not report on standard groups because they truly do not use them or because they simply failed to use that segment of the log.

The teachers' response regarding "standard" groups also raises a conceptual issue: when is a grouping standard and how often can or should the teacher update those group lists? Without a clear definition of standard groups, and a method for changing their makeup, we may find that teachers who use groups do not report them as standard. There could also be social reasons not to acknowledge using groups, especially if the groups are essentially tracks or ability groups, which are discouraged in some schools and districts.

Student activity

Beyond what teachers reported about individual students, the log also captured information about activities that were used to engage students in these tasks and the ways in which teachers organized students for instruction. We asked teachers whether they taught the class as a whole, whether the class was broken down into small groups, or whether students engaged in independent work during the lesson.[9] Since teachers were reporting on an entire lesson, selecting any combination of these three grouping formats was permitted. In well over half of all lessons (62 percent), teachers said they utilized "whole group lecture or discussion." This format was considerably more prevalent in math lessons. In roughly half of all lessons, students were placed into small groups. This form of organizing students occurred more often in reading lessons. Roughly 40 percent of all lessons involved students in independent work. Such work was more likely to occur in reading lessons. Since we allowed teachers to select multiple grouping patterns for each lesson, it was possible to examine the different combinations of formats used by teachers. Slightly more than half of all lessons utilized only a single form of grouping. Among multiple format lessons, the combination of whole group and small group was most prevalent, with about 20 percent of all lessons utilizing both these formats.

Beyond documenting the ways in which they were organized, the log attempted to capture what students actually did during the lesson. For example, when teachers reported on their reading lessons, they answered a series of questions about the nature of student activity in the lesson. Teachers were first presented with a list of seven possible choices for student activity including: reading with trade books, working on textbooks, working on worksheets, working with flash cards, working on a computer, playing instructional games, and working on a teacher led activity. If student activity fit none of these categories, teachers could indicate "other activity." In a majority of the reading lessons, teachers said that only one student activity was utilized. In about 40 percent of the lessons, teachers reported two different student activities. In a handful of lessons, teachers reported that students engaged in three different activities.

The various categories of student activity were chosen at roughly equal levels of incidence. Exceptions to this pattern were "working in textbooks" and "working on a computer." Teachers chose each of these categories only once. The small number of cases in the pilot makes it difficult to judge whether computers and textbooks are useful categories for teachers or whether they are just more rarely used. Two other results are noteworthy. The first is that "other" represented one-quarter of all student activity. Engaging in discussion, writing, and answering teachers' questions are commonly mentioned student activities that do not appear on our list. The heavy use of the other activity category suggests that our list of student activities may be not comprehensive, poorly defined, or not understood by teachers.

Teacher activity

The log prompted teachers to record information about adults who delivered instruction during each lesson. Choices ranged from indicating the presence of multiple teachers to recording lessons for which the regular classroom teacher was not in the classroom (e.g., when a substitute taught the lesson).[10] When there was a paraprofessional or other adult teacher in the room, the log asked for information on his/her instructional role as well as the classroom teacher's role. In the pilot data, teachers reported a paraprofessional teaching during the lessons four times and another adult teaching during the lesson five times. Out of the 29 lessons, teachers reported four lessons that were taught by two teachers. Three of these lessons were in reading and one was in mathematics.

The teacher log also provides information on teacher activity. This version of the log asked teachers to report on the following activities:

- Worked with individual children
- Worked with students in a small group
- Circulated around the room
- Lectured or led discussion
- Non-instructional activities

Out of the 29 lessons recorded, teachers reported working with individual students in 38% of the lessons, working with small groups of children in 41% of the lessons, circulating the room in 72% of the lessons, lecturing or leading discussion in 63% of the lessons, and doing non-instructional activities in 3% of the lessons. For many lessons, teachers reported more than one teacher activity. The frequency of teacher activities breaks down in the following manner: teachers reported doing only one teacher activity 31% of the time, doing two teacher activities 34% of the time, doing three teacher activities during one lesson 22% of the time, and doing four teacher activities 13% of the time. During math lessons, teachers most frequently reported doing two or three teacher activities (10 lessons out of 14), while in reading instruction teachers most frequently reported doing one or two teacher activities (12 lessons out of 15).

Topics and materials

In addition to collecting information that allows us to map differences in the way teachers delivered instruction and students engaged in it, we sought to describe the content of each lesson. In this version of the log we collected content data in three ways. First, teachers listed topics for each lesson. Second,

teachers reported on how the content of the lesson was conveyed to students, what we called "nature of the task." Finally, teachers listed the materials with which students worked.

For the log pilot, teachers were asked to enter the topic(s) for each lesson in a text field. No topic list was provided by the system, and teachers could enter more than one topic for each lesson.[11,12] The total number of topics teachers reported for their four lessons ranged from a minimum of four to a maximum of nine. The teacher who recorded eight lessons reported thirteen topics. The other six teachers reported four, four, six, seven, eight, and nine topics respectively. In general, teachers reported one topic per lesson. In cases with an average of two or more topics per entry, the high average was often due to one or two particular lessons in which multiple topics were taught.

There is also a reporting trend toward fewer topics per lesson near the end of the research pilot. It is possible that the instrument led teachers to report only one topic per lesson. The log pilot is designed in such a way that teachers could choose to list a few topics (e.g., main character, setting, and plot) together as one topic or on separate lines as separate topics. Teachers had a strong disincentive to list such items separately since the log then prompted them for additional information about each item. A number of teachers mentioned that they ended up choosing to list related content topics together rather than apart. The data supports these comments since the majority of multiple-topic lessons were entered near the beginning of the log pilot.

For each topic entry, teachers were linked to a screen which asked them to report on the nature of the task in which the students were engaged. The number of reports on nature of the task should equal the number of topics. We found that the items concerning nature of the task were often skipped. In this version of the log, topic entries were created when the teacher typed a period or semicolon or hit the return key in the text field. When teachers did these things without intending to have more than one topic, the log presented the teachers with duplicate 'nature of task' screens and may have led teachers to ignore one screen.

The log also recorded, for each topic, the teacher's evaluation of student engagement and the purpose of the task. In this pilot teachers never reported student engagement as low. They did report a fairly even split between moderate and high. As a group, teachers reported moderate engagement for 21 of the 54 topics they reported and high engagement for 26 of the 54 topics. They did not answer for seven of the topics. Finally, teachers were prompted to indicate whether the lesson was intended to introduce, further develop, review, or conclude the specified topic. Teachers did not answer the question about the purpose of the topic for seven of the topics. The most frequent teacher response to this portion of the log was 'further develop'. This response was given for 26 of the topics. The rest of the topics were split evenly among the other choices (8, 7, and 6 responses for introduce, review, and conclude, respectively). This is interesting given the fact that the pilot was conducted during the last three weeks of the school year. Will the distribution of purposes look different at other times of year?

Results from our initial analysis show that teachers almost always reported about the materials they used in their math lessons. In 13 of the 14 math lessons, teachers reported the text they were teaching from and (in all but two of the math lessons) the publisher. In addition, log data indicated that the major source of math materials was a course or workshop (recorded for 13 of the lessons), the materials were usually from a published text or series (in 12 lessons), and these materials were generally required by the district (in 11 of the lessons).

The data we collected about materials did not give any sense of how the teachers used the materials. Specifically, we are still working on ways we might learn how closely each teacher follows a text or curriculum guide, whether she/he uses the materials as intended by the authors, adapts them to her own purposes, omits or supplements the content, or alters the approach suggested by these materials.[13]

The Validity and Reliability of Teachers' Responses to the Log

A validity analysis was conducted using the 24 lessons for which we obtained reports from both teacher and researcher. Four of these lessons included two researcher reports. This raised the question of whether to include both researchers in the count of agreements or disagreements between researcher and teacher. For this analysis, we randomly chose data from one of the researchers to compare to the teacher entries in each of these four lessons.

We analyzed by item data at the lesson level and data which could be aggregated to the lesson level. For items which permitted multiple responses (e.g., student activity in math), dummy binary variables were created for each response. That is, the teacher or researcher either marked or did not mark a given student activity (e.g., worked on worksheets) for a given lesson. If both the teacher and researcher marked an activity as included in the lessonæor if they both reported it as absent from the lessonæa case of agreement was recorded. A case of disagreement was counted when teacher and researcher responses differed for the same lesson.

Data with lesson-level information were compared by lesson using line graphs and scatter plots. The line graph highlighted differences by lesson between teachers and researchers. Scatter plots illustrated how frequently the cases of disagreement resulted from a teacher response and researcher non-response, or vice-versa. Data with multiple cases per lesson (e.g., topics) which could not be aggregated to the lesson level were analyzed differently. First, we analyzed whether the teachers and researchers broke the lesson into the same number of topics. Then, we aggregated the teacher and researcher responses to the following variables: the nature of the task students were working on, the student's engagement, and the students' prior exposure to the material. Teacher and researcher data were then compared.

Most items in the teacher log were collected at the lesson level. Some of these items are general and apply to either subject (math or reading) taught by the teacher. Other items are subject specific and were looked at separately. The lesson-level items analyzed below include four generic categories of information: duration of the lesson in minutes, who was teaching the lesson, what activities the teacher was doing during the lesson, and the student grouping pattern for the lesson. Subsequent sub-sections summarize additional, subject-specific information.

Agreements on lesson duration

In nine of the twenty-four lessons, teachers and researchers reported the same amount of time for the lesson. Therefore, in almost forty percent of the lessons teachers and researchers matched each other exactly on the amount of time for the lesson. However, the data on the whole are not very encouraging. On average, teachers and researchers differed in their reports of time by over 8 minutes per lesson. On average, lessons lasted somewhere between 46 and 54 minutes. This range represents the average of the lowest reports in minutes and the highest reports in minutes. Using the high end of the range, this

means that researchers and teachers differed one minute for every six minutes of lesson. This ratio is surprisingly high for such a low-inference item.

The following are other ways to characterize differences in teacher and researcher reports of duration. Teachers and researchers differed by approximately eight minutes on average. For fourteen of the twenty-four lessons, teacher and researcher reports of the duration differed by eight minutes or less. This works out to just over 58% of the lessons. Using ten minutes as the bar, teachers and researchers differed by ten minutes or less in 79% of the lessons. Even if we accept that teachers and researchers could be off by as much as ten minutes, there are still 21% of the lessons where teachers and researchers differed by more than ten minutes in lessons that lasted, on average, less than an hour.

There are many possible explanations for these differences in reports of time. Although researchers arranged with teachers to observe for a "lesson," the definition of a lesson, and what constitutes the beginning and end of a lesson, was left up to the teacher and researcher. Unless both the teacher and researcher agreed upon a definition, either of the two might have reported the amount of time the researcher was present in the classroom. In addition, portions of the lesson may have been considered "instructional" and part of a lesson by one of the two, but not seen as "instructional" and *not* part of the lesson by the other. This is informative for future work. It highlights the difficulty in determining what is and is not part of a lesson and, therefore, points to the necessity of carefully defining what constitutes a lesson or other structure for recording instruction.

Agreements on who was teaching the lesson

In the log, teachers were asked to report on who was teaching the lesson. In addition to the classroom teacher, teachers were given the following options: a paraprofessional, a parent, an intervention teacher, or other. Dummy variables were created for who taught the lesson, whether the teacher or researcher marked classroom teacher, paraprofessional, or other. In no cases did the teacher or researcher report involvement of an intervention teacher or parent in teaching a lesson; therefore, these variables are not included in this analysis.

In twenty-one of the twenty-four lessons (88% of the lessons), teachers and researchers were in agreement over whether or not the classroom teacher taught the lesson. For twenty of these lessons, teachers and researchers both reported the classroom teacher as teaching the lesson. The three cases in disagreement occurred with the same teacher-researcher pair. In this case, the teacher is a floater/professional developer at her school. She frequently teaches in different classrooms and mentors many teachers. However, since all of the lessons for the pilot occurred in the same classroom, the researcher considered her the classroom teacher during the pilot.[14] Since all of the cases of disagreement occurred as a result of this situation, the 88% agreement is seen as a conservative estimate of the potential validity of this item.

There were only three lessons in which either the teacher or the researcher, or both, reported a paraprofessional teaching the lesson. In one lesson, both teacher and researcher reported the paraprofessional as a teacher. In two other cases, either the teacher or the researcher reported a paraprofessional teaching the lesson and the other did not. This yields twenty-two cases of agreement and only two cases of disagreement. Again, the 92% agreement may actually be a conservative figure. It appears that for one lesson either the teacher or researcher reported there being a paraprofessional

teaching the lesson while the other person reported that there was an "other" teaching the same lesson. This issue may be addressed in future versions by more carefully defining our use of the term paraprofessional.

Finally, the use of "other" as a teacher during the lesson was limited to only five lessons. The aforementioned teacher reported three of these cases. However, only one of these cases appears as a disagreement by this teacher-researcher pair. There are two other disagreements, for a total of three. This means, similar to the classroom teacher item, there is 88% agreement for this item.

Agreements on what the teacher was doing

Five dummy variables were created for teachers' responses to the following question in the log: "During lesson, the teacher: worked with individual children, worked with small groups, circulated around the room, lectured or led discussion, [or] non-instructional." Teachers were told they could respond to more than one of these items, if appropriate. Across the 120 opportunities for teacher-researcher agreement (24 lessons, 5 teacher activities per lesson), there were 94 cases of agreement reported between teachers and researchers, a 78% agreement rate. This rate is likely to be improved upon by better defining the categories and by defining criteria to guide teachers' decisions to mark, or not mark, an activity. For example, if a teacher circulated the room for less than five minutes during a one-hour lesson, he/she may have wondered whether or not to mark that option. Establishing clear criteria for this question should take the guesswork out of teachers' responses. We may also decide to ask teachers to record which activity predominated in each lesson.

Although there is only slight variation in teacher-researcher agreement rates in the five categories, there appears to be greater disagreement in categories for which interpretation of the item came into play. For example, the item with least agreement (17 out of 24) asked whether or not the teacher worked with individual children. In this pilot, no guidelines were given as to what constitutes "working with an individual." Therefore, one person may have considered it to be five minutes of sustained instruction with one student, while another person may have considered three minutes of sustained instruction with a pair to constitute work with individual students (plural). Similarly, over the course of the pilot, questions arose regarding what constitutes a small group. Therefore, it is not surprising that this item had agreement in only 75% of the cases. This fact is not meant to diminish the importance of this number. The fact that we received 75% agreement without providing explicit instructions should be viewed as encouraging.

The disagreements which arose in this section of the log are most likely due to the fact that teachers and researchers could report multiple activities as part of any lesson. Teachers and researchers only reported the same number of teacher activities for the lesson in eleven cases (46% agreement). Even in these cases, there is no assurance that teachers and researchers recorded the same teacher activities. Therefore, the validity of these items may be improved by defining the categories more precisely and by establishing criteria for including teacher activities in the log.

Agreements on student grouping patterns

Teachers were asked about the grouping pattern for the lesson and given the following three options: whole group, small group, or independent work. Teachers could choose more than one option if applicable to the lesson. The greatest teacher-researcher agreement rate occurred in reports of the

presence or absence of whole group instruction. This may be due, in part, to the fact that only seven lessons did not include whole group instruction. In three of these lessons, both teachers and researchers reported no whole group instruction, while in the other four, researchers reported whole group instruction and teachers did not. Teachers and researchers were in agreement for twenty of the twenty-four lessons.

The responses on small group work and independent work were similar, and they differed from the responses on whole group work. There was slightly less agreement between teachers and researchers for these: 75% for small group work and 71% for independent work. There was also a more even distribution of responses across the 24 lessons. Teachers and researchers agreed on eighteen cases. In nine of these cases, they agreed that small group work was present in the lesson; in nine they agreed it was absent from the lesson. Teachers identified small group work in three lessons and none in the remaining three; researchers made the opposite identification for each of these six lessons. Similarly, for independent work teachers and researchers agreed that it was present in eight lessons and absent from nine lessons. However, there were five lessons in which researchers, but not teachers, felt there was independent work. There was only one instance in which a teacher marked that option when the observing researcher did not.

A probable reason for the higher agreement rate of whole group work is the fact that it was present in nearly every lesson. However, lack of clarity in the log choices probably contributed to the disagreements in this section. The definition of "whole group" is obvious, while the definition of "small group" is left to the interpretation of the teacher or researcher.

Agreements in analyses of reading lessons

This pilot included different branching schemes for math and reading lessons. Therefore, questions about student activity and materials used by students were asked differently, and the data was stored in different databases. For this reason they were also analyzed separately. The question about student activity in reading contained eight options (students read trade books, worked in textbooks, did worksheets, worked on computers, worked with flash cards, played instructional games, did a teacher-led activity, or did an "other" activity). Teachers could respond to more than one option. Once again, dummy variables were created for each of the options in order to analyze the percent agreement between teachers and researchers.

In no cases did teachers or researchers report either "worked with flash cards" or "played instructional games." Although these technically count as an agreement between teacher and researcher, they have been left out of this analysis to arrive at a conservative estimate of validity. The sum of the agreements of the other six items resulted in an agreement rate of 75%, or 54 agreements across the 72 opportunities for agreement in the reading lessons.

This number is encouraging considering the fact that "teacher-led activity" and "other activity" were not well defined in this pilot. These two categories comprise half of the disagreements between teachers and researchers. Furthermore, the lack of definition given to these categories may account for differences elsewhere. Teachers who used worksheets but who led students through the worksheet may have marked "teacher-led activity" rather than "worksheets," or they may have marked both categories.

Here, as in other parts of the log, it is clear that we need to better define our categories and provide more direction on how teachers are to characterize certain activities.

It was difficult to analyze "reading materials used" due to the way the database was designed. For each reading activity, teachers were asked to enter the text and publisher used for the activity and to indicate who selected the materials for the lesson, whose decision it was to use the materials, what the source of the materials was, and, if students used trade books, whether or not students selected their own trade books. When aggregated to the lesson level, the sum of certain items (e.g., those indicating whether or not a text was used for reading) was often more than one. This means that teachers and especially researchers may have input the text used for the lesson more than once, *or* the lesson may have required entering more than one text. In this analysis, any number greater than zero was interpreted as a "text used" entry made by the teacher. This problem of redundancy has been commented on previously, and we have suggested several ways of dealing with it. Perhaps the easiest solution is to create a pull-down menu with a default yearlong text or list of texts for teachers to access when completing the log.

Agreements in analyses of mathematics lessons

The analyses of the math lessons followed the same procedure as that described for the reading lessons above. However, the questions asked in the log were slightly different for reading and math lessons. For math lessons, the "student activity" question asked teachers to characterize the materials with which students were working. The options given teachers included: manipulatives, worksheets, flash cards, calculators, instructional games, teacher-led activity, or other activity. None of the teachers or researchers indicated the use of flash cards in their lessons, so that option has been omitted from this analysis to yield a conservative estimate. The sum of the agreements across all student activity categories is 53 (not including flash cards) out of a possible 72 (12 lessons by 6 possible categories per lesson). This results in an agreement rate of 73%.

We also compared the number of student activities supplied by teachers to the number marked by researchers for the same lesson. This item had low validity. Further, even when teachers and researchers reported the same number of student activities, we had no idea if they are the *same* activities. This result suggest that we need to define what does and does not constitute a student activity. For example, do students need to engage in a particular activity for a given amount of time for it to "count" as an activity? Some guidelines and definitions of our items should enhance the validity of this data.

The data collected on materials for math lessons was encouraging. For every math lesson, there was a text entered by both the teacher and the researcher. Further, looking at the actual string data, the text reported by both teacher and researcher clearly matches in all cases but one. In eight of the lessons teachers and researchers also input the publisher associated with the text. Teachers entered the publisher in three additional lessons when the researcher did not, for a total of 11 publisher entries for 12 lessons. This suggests that information on materials, especially with more sophisticated means of data collection (e.g., pull-down menus), should be accurate and moderately easy to retrieve. In cases in which a teacher entered the publisher but the researcher did not, the researcher may not have had information about the publisher. Another possibility is that after entering the publisher information once, the researcher did not continue to do so in the following lessons.

Finally, data about the source of the materials, the type of materials, and the decision to use the materials all indicate high levels of agreement between teachers and researchers. Compared with many other variables, these three variables are all three-level categorical variables. When viewed in this context, the high rates of agreement for these variables are even more impressive. For example, in the "decision to use" category, both teachers and researchers reported that the district requires the use of the materials for six of the lessons, and they agreed that the teacher selected the materials for another three of the lessons. Similarly, there was a high level of agreement about the source and type of materials. This can likely be attributed to the fact that researchers had to inquire about the materials used in lessons they observed in order to input their data in the log.

Reliability of log data

Little prior experience is available to inform us about the reliability of self-reported data on instructional practice. We recognize problems related to the lack of a shared professional lexicon for content topics and instructional approaches. Although researchers did observe the lessons logged by the teachers in the pilot, we witnessed too little to estimate reliability of teachers' data. Both discrepancy and alignment were evident, but we were not able to see patterns in either. The pilot was not especially useful in advancing our understanding of the reliability of teachers' self-reports.

Teachers' Willingness to Use the Log

To understand how willing teachers might be to use the log, we paid attention to the sense of burden they might experience in filling out the log and to possible sources of incentive for them to do so. Because this pilot was of such short duration and the teachers involved were selected in part because of their desire to participate, we did not get much useful information about the burden that might be felt by teachers in the study. What we did learn is that using a Web-based tool is feasible for teachers who have Web connections in their schools. Teachers commented that they would like to be able to use the instrument at home. (This was possible during the pilot but was not actually tried by any of the teachers.) We saw some evidence that teachers began to economize their input even in this short pilot. For example, they consolidated multiple topics into one topic, limiting the number of screens they were asked to complete (see the section above on "Topics and Materials").

During the focus group meeting at the end of log data collection, we asked the teachers to share their ideas on ways to improve the log and to suggest reasonable incentives to encourage teachers to use the log in the future. These two areas of discussion melded into one, as teachers focused on what would make the log useful to them. Their suggestions for improvements generally fell into three categories: those that would give teachers more information about their students, those that would help teachers plan or reflect upon their practice, and those that would assist teachers with administrative tasks or reporting requirements. Suggestions for other external incentives included financial support and release time.

Our discussions with teachers suggested that log enhancements have the potential to serve as an incentive for filling out the log. All of the enhancements mentioned below were things that teachers identified as being useful and informative to them. It seems logical to expect that the more the log provides teachers with useful information about themselves and their students, the more likely they will be to use it. However, it is important to keep in mind the context within which teachers made these

suggestions. They only had to complete the log four times. This falls far short of the time commitment we will ask of teachers in the actual study. Also, in our focus group, we essentially asked teachers to imagine enhancements they would want without paying particular attention to possible increases in burden that might result from such enhancements. Clearly, any of the enhancements discussed below would require a greater time commitment from teachers and might create a greater sense of burden. As we test possible enhancements, the tradeoffs between potential benefits must be examined alongside potential detriments (e.g., through increased burden). In addition, we must consider the responsibility for data integrity and security that the project would take on if we kept any kind of official records for teachers, as well as the programming and development burden of actually implementing complex student record management across multiple schools and districts.

Enhancements that give teachers information about their students

Teachers told us that the log would be much more useful to them if it yielded usable information about their students. The most widely cited enhancement suggested by teachers was having a means to track the progress of individual students over time. A number of teachers felt it would be useful to create a record of the dates on which individual students understood specific topics and of the amount of effort required by the teacher to achieve that understanding. All of the teachers in the pilot came from a district which is moving to a complex assessment requiring them to track student progress on a number of objectives for each subject. They saw the possibility of using an instrument like the log to track student progress once a day rather than taking time from classroom instruction to meet with each student for individual assessment. Another teacher said that he would use information on individual student understanding in a diagnostic fashion to help guide his decisions about which individual students he needed to work with on particular topics.

At a more basic level, some teachers thought that having a historical record of which students were absent when particular topics were covered would also provide a useful diagnostic tool. They felt that such information might help them understand why students were struggling with particular topics and subsequently might help them shape instruction in response to the struggles.

In addition to being able to track information on individual students over time, a few teachers thought it would be useful to record the progress and activities of groups of students over time (e.g., reading groups).

Enhancements that help teachers plan or reflect upon their practice

As the previous section suggested, enhancements which provide more student information to teachers are considered desirable because they would help teachers shape their practice in response to their students' needs. Teachers noted a number of other enhancements related to their' practice. A few teachers expressed interest in using the log as a planning tool for their teaching. One teacher who spoke on this issue acknowledged that the lesson plans he filled out and saved did not contain very much useful information. He suggested that lesson plans entered into the log might be more useful than traditional ones because they would be entered alongside descriptions of what actually happened in the classroom, thus setting teachers' plans within the context of actual classroom practice.

A number of teachers mentioned that they would like a historical record of how much time they spent on specific topics over time. For some, the utility of this information resided in the opportunities

it provided for teachers to know which students were exposed to which topics and subsequently which students understood which topics.

In addition to historical information about themselves, some teachers thought their teaching might be enhanced if the log helped them connect to information from professional communities of practitioners beyond the schools in which they work. One teacher viewed working on a study being conducted by a school of education at a major university as a plus because of the information about teaching she might gain through her participation. This teacher was very interested in tapping into new ideas about practice that might inform her teaching.

Enhancements that assist teachers with administrative tasks or reporting requirements

As discussed above, a number of teachers said that being able to record students' understanding of specific topics would be a great help to them in their greatest reporting task: the "K-2 Assessment." The teachers felt that clicking on which students "got" a topic during class would be much more efficient than the current practice of pulling each student out of class. In a similar vein, a few teachers felt that having a log module in which student grades could be entered would be very helpful.

Other incentives

Teachers mentioned financial incentives, but they were more enthusiastic about receiving various types of classroom assistance—such as someone to help with record-keeping or grading assignments—in order to give them more time to complete the log.

Technical Results

Technically the log worked quite well. After a brief orientation in which researchers helped teachers enter class and group lists and work through a trial log entry, teachers were able to access the log and enter data without encountering severe technical problems. None of the teachers reported losing data from log entries. Features such as the ability to call up reports of past lessons also worked very smoothly. Data from teacher entries was successfully captured in a database and subsequently prepared for analysis.

Database structure

The data structure we developed in FileMaker Pro was not adequate for the kinds of data analysis we wanted to perform. Because the amount of data was small, we were able to reconfigure information into more usable formats. However, in the actual study, we need to develop data structures, which are congruent with the way the data will be analyzed. Several important issues were apparent:

- *Unique identifiers for students, teachers, and lessons.* Some of these identifiers will need to be coordinated with other data in the study. Within the log database, the lesson identifier is of particular importance because of the likelihood of multiple reports on the same lesson (from observers in the classroom, from other classroom teachers whose students move across classrooms, or from other teachers working in the same classroom).
- *Level of detail and units of analysis.* Some data will be collected at the individual student level, some at the lesson level, and some at the classroom or teacher level. The pilot identified the need to make decisions about which data belongs at which level and how each type of data will be used in analysis.

- *Level of detail collected for individual students.* It is still not clear exactly what "picture" of instruction we want, or can collect, at the student level. However, the database needs to be carefully designed to make it possible to associate data by student whenever possible.

Web page design

Web page design was quite successful for the pilot. Teachers reported that they had little trouble figuring out how to navigate within the log, and we saw few problems with the design other than the conceptual problems already mentioned. As we had hoped, the log successfully provided branching based on teacher input and presented customized lists of students for each teacher. One problem, which surfaced as the researchers, continued their own analysis and discussion of the pages was that the layout was slightly different for reading and mathematics. This led to confusion in interpretation and analysis of the data and possibly some anomalies in the way the pages were filled out by teachers. In general, it seems clear that fewer page changes are better and that a consistent path through the pages for reading and mathematics is desirable.

Web pages provided as "help" for teachers in understanding the meaning of log questions and in navigating through the log were not used by any of the teachers in the pilot (see the section on Training below for more discussion of this).

Training

For the pilot, researchers worked with teachers to create dummy lessons in order to illustrate how to use the log. In some cases, the researcher actually entered the class list; in other cases, the researcher helped the teacher while she typed in student names. During the focus group discussions, teachers indicated that they wanted more training than they received before using the log. In future pilots, not only do we need to clarify terms, but also to provide adequate training for teachers in order to be certain that the log terms are interpreted and used correctly.

Although extensive help documentation was provided both on-line and in print, we saw no teachers use either document, and teachers reported that they had not read any of the documentation. This presents a problem for training: how can we communicate our definitions for key terminology and distinctions to teachers in a way that will get their attention and that will work consistently for all teachers involved in the large study? However, in light of the brevity of the training we provided for the pilot and the teacher's inattention to the documentation provided, it is encouraging that teachers were both able to understand and use the technology of the log and achieve an acceptable degree of validity for many items.

Deployment

Although we had few problems with deployment of this log, the problems we encountered served as reminders of the complexity of implementing software and hardware systems. First, there were numerous glitches in the database itself which had to be resolved by the programmer. These were things such as having the input mechanism on the server overload and need to be reset or having a table in the database close unexpectedly and need to be restarted. None of these resulted in data loss, but in a larger scale deployment, a higher level of reliability will be essential.

Second, we had some server problems, such as unexpected crashes. Again, we were able to resolve these without loss of data, but only because the scale and demands on the server were manageable.

Finally, at the schools, we had to go to each classroom and install the correct version of Netscape because our Web pages could not be viewed adequately with the district standard version. (The district used version 2.2, and we needed to have version 3.0 or higher to view tables.) The computers being used by these teachers were in some cases older model Macintoshes which were barely able to handle the newer version of Netscape. In addition, on one computer, installing Netscape led to other problems which ultimately caused major problems on the hard drive requiring complete reconfiguration of the machine and an unfortunate loss of some of the teachers personal data. All of this points to the need to establish a minimal configuration for use of the log. In particular, we will establish and require minimum browser versions (for Netscape and Internet Explorer), minimum memory and hard drive sizes, and minimum line speeds. Our software will be kept as simple as possible, with no requirement for Java capability or other special add-ins. Although this limits what we can do, it seems essential to minimize the customization that might be needed on teachers' computers.

Discussion

We are encouraged by a number of findings from the pilot study. One purpose of this work was to test the practical and technical feasibility of a Web-based instrument. The development of a Web-based log was a complicated undertaking that entailed developing and implementing instruments largely untested in social science research. Just the fact that we were able to implement a Web-based instrument that successfully allowed teachers to record classroom instruction and yielded data suitable for preliminary analysis is encouraging.

When we began this work, some researchers we consulted questioned whether teachers would be willing to interact with a complicated and technical instrument. Although our sample was both small and non-representative, we are encouraged by the high level of commitment and interest demonstrated by the teachers. Without exception we found that teachers were thoughtful, eager to participate, and had a professional interest in the data we were collecting. This work suggests that instructional instruments that are designed to capture and represent data about instruction in ways that provide teachers with records of their practice may also afford powerful incentives for teachers' involvement.

We were also uncertain, when we started this work, if we would be able to use Web-based technology to develop the kind of computer instrument we envisioned. The pilot demonstrates that Web-based technology is suitable for a sophisticated branching instrument, that the Web can be used to capture instructional data directly into a database, and that it is possible to develop instruments that teachers can use and understand with reasonable levels of training and support.

Based on our analysis of the pilot data, we think a Web-based instrument could be used successfully to capture some important information about teaching, some of which has not been available before. We are encouraged about the possibility of capturing the following:

- Number of reading and mathematics lessons
- Duration of each lesson
- Student attendance for each lesson
- Materials used for each lesson
- Instructional groupings or formats for each lesson

These items would be valuable in understanding instruction and in evaluating the instructional resources available to each student in the study. Even if the log were able to capture just this much, it would provide useful data about instruction and instructional interventions.

Other items would enhance the picture of instruction, but they need more work in order to present clear and unambiguous choices to teachers. These items are more inferential, and some have not been studied through teacher self-reports:

- Topic: what content is presented to the students?
- Nature of the task: how is the content presented to the students?
- Student activity: what do the students do during the lesson or with the content?
- Teacher activity: what does the teacher do during the lesson?
- Purpose of the lesson.
- Student engagement with the lesson.

Outstanding Issues

Whatever its successes to date, the Web-based log we have developed raises important issues of validity, reliability, sampling, burden, and teachers' interaction with the instrument. These will be high priority challenges for future development work.

Validity issues

We need to have confidence about what the items mean to teachers. Central here is understanding how teachers interpret items and how they represent their teaching through their responses in the log. Thus, there are threats to validity that deal with teachers' understanding of the log's questions, language, and conceptualization of instruction. Other researchers have found that the validity of items related to instruction is weaker with respect to instructional practices most associated with current reforms (e.g., everyone is likely to report that they hold discussions in class, engage children in literature, or teach problem solving). Another potential problem entails teachers' ability to recall details of a lesson many hours after that lesson occurred.

Reliability issues

We also need to design our next pilot to probe issues of reliability in teachers' reports. Among the crucial issues are whether teachers will report similar elements of instruction in consistent ways across time and whether specific elements on the log will be interpreted similarly across the group of teachers. In addition, we need to ascertain the reliability of less inferential data such as time allocations to subject matter, student attendance, and materials usage.

Sampling issues

We need to examine closely which aspects of practice can be sampled (e.g., data on the whole school day, grouping data, attendance, materials used, time allocated to math and reading) given our questions and planned analyses. We have a provisional list of variables for which we want to collect information about teaching. A big question for us is how much of a teacher's practice we must seek to record in order to have reliable and valid data to conduct the analyses we want. Several issues complicate this question.

First, we know little about how particular instructional variables are distributed across a teacher's practice. No research has attempted to answer this question. In fact, virtually all research on teaching has, by default, sampled teachers' practice and has made inferences about practice that assume that such samples afford accurate portraits both of individual teachers' instruction and teachers' practice more generally. What sorts of regularities exist in teaching and what is variable? For example, do teachers consistently emphasize the same learning goals or do these vary depending on the topic, the students, the time of day, week, or place within the unit? Is the amount of time spent on mathematics or language arts stable across a given time period (e.g., over a whole instructional unit, month, or year)? Or, are there periods in which much more time is devoted to one of these subjects while little time is allocated to the other? We need some way to gather evidence that would yield some insight into these potential sources of variation.

Second, some elements of practice cannot be reasonably sampled. If we want to collect data on content coverage and the time spent on particular topics, for instance, we likely need to collect information every day. Similarly, in order to ascertain which students received what instruction, we need daily attendance and tardiness records. One question, then, is the extent to which we will need comprehensive information about content coverage and attendance in order to estimate relationships between opportunity to learn and student performance. How much of our analysis is likely to seek close mappings of particular topics with specific student competencies?

Third, although the first two issues might lead us to want to collect log data every day, practical considerations discourage this. No matter what we do to design the log sympathetically, it will ask a great deal of busy teachers. The more days on which we ask them to fill out the log, the greater the burden we are likely creating. One possibility is to create a "light" version of the log that could be administered every day, asking only whether reading or mathematics was taught; if so, what topics were addressed; and who was present in class. Other instructional variables would be measured on some other time sampling plan.

Issues about scaling up

The pilot study was conduced with a small, select sample of teachers over a short period of time. We were able to work closely with teachers as colleagues, listening carefully to their ideas and suggestions. Teachers were included as part of the research team as we worked through problems and incorporated suggestions. This level of personal contact would be hard if not impossible to replicate with a much larger sample of teachers. In order to collect data that usefully informed questions about instructional improvement, the log would need to work in a much larger sample of classrooms, with a larger variety of teachers, over a much longer period of time. We still know little about what it mean to implement the log with a large number of teachers of teachers from diverse backgrounds over long periods of time. However it is reasonable to expect that scaling up an instrument such as the pilot log would introduce problems with keeping teacher data entries up to date, maintaining the quality of the data, and keeping the involvement of the teachers.[15]

We also still have a great deal to learn about how much teachers will be willing to do. What it will take to get teachers to complete the log regularly, whatever the time sampling plan, is something for which we still have little information. Are we being naïve, imagining that 900 teachers will actually consistently fill out logs on their reading and mathematics lessons over a relatively long period of time?

Are there other issues—such as teachers trusting that the data collected will not be used against them—, which will interfere with data collection? Alternatively, are we perhaps overestimating the burden of this task?

Issues about teacher interaction with the instrument

Teachers in the pilot expressed great interest in using something like the log to help them in their own practice, both for record keeping and for tracking students in ways that have not previously been possible. However, this possibility raises issues about the nature of the instrument we are creating: to what extent can it be an agent of change itself before it loses its value as a data collection instrument? The pilot work, however, does point to the promise of a number of modifications to the log which could help teachers learn to use the instrument and could provide them feedback about their own teaching. These include:

- *Providing on-line help, which explains log items to teachers when they first begin to use the instrument, but which disappears as they become familiar with the log.* For example, during the first weeks of use, we might include complete explanations and examples for each category in "nature of the task", but after some period of time these explanations would either cease to appear or the teacher could choose to hide them. We consider this because of the well-known difficulty in getting software users to consult extraneous documents (manuals or even hidden on-line help) as they learn to use a program and because of our pilot experience, during which no teachers used our documentation to learn to use the log.
- *Providing reports to teachers about their previous lesson reports or about student attendance.*
- *Leaving lesson reports open for a number of days after their initial creation to allow teachers to modify their report.* We consider this because of the press of time: it might lead to better reporting if teachers had more time to refine their reports. In particular, they might need to start a report for a given day, stop in mid-report, and pick it up later; or they might fill in a report in haste and want to improve their data at a later time.

Each of these options has implications for validity and for the impact of the instrument itself on each teacher's practice. They also have implications for what we can learn about the details of instruction in the interventions.

EPILOGUE

Following the initial pilot of the Web-based log which we have described in the preceding pages, the study team proceeded with further developmental work and pilot testing. Our goal in this later stage of developmental work was to see if efficient and technically sound ways could be devised for handling the various issues we noted above.

Although we were able to make further improvements in the Web-based instrument, we decided that it did not seem feasible given the scope of our data collection. We will be collecting data in over a hundred schools and in a variety of environments, and we realized that the scale and variability made reliance on the web unfeasible. We continued with the development of a paper and pencil version of the instrument. This decision reflects emerging study-specific constraints and in no way diminishes the potential viability of a Web-based instrument for a variety of other research purposes.

ENDNOTES

1 The interventions are familiar, including those in the New American School Designs and several others that are at least equally promising. The Obey-Porter Amendment passed by Congress within the last year is only the most recent expression of increasing efforts to re-focus schooling and school improvement on results for students. Both of these developments occur against a background of equivocal or negative evidence on the effects of school improvement projects and programs, especially in high poverty schools, and increasing concern that Title I may be in a fight for its life in the 1999 reauthorization.

2 However, given the unpredictable policy environment surrounding schools, it is possible that during the course of the study some of these schools will launch their own school improvement initiatives. This, too, will be of interest to us since it allows for a more dynamic scenario in which to compare and contrast school improvement efforts.

3 See Knapp, M. S., & Associates (1995). *Teaching for meaning in high-poverty classrooms.* New York: Teachers College Press. The log instruments and procedures used in this study are more fully described in Knapp, M. S., & Marder, C. (1992). *Academic challenge for the children of poverty—Volume 2: Study design and technical notes.* Washington, DC: U.S. Department of Education, Planning and Evaluation Service.

4 Since these data will be used in part to understand how instruction influences reading performance, topic domains in reading and mathematics need to be valid in reference to both the assessments used in the larger study and the instructional design of the interventions we will study.

5 Andrew Porter and his colleagues pioneered the use of logs in studying teaching in their groundbreaking work on the content determinants of elementary mathematics teaching. Their logs were paper-pencil documents that collected information on what teachers were teaching. Using interviews, the researchers probed the content determinants of these curriculum records. See for example:

Freeman, D. J., & Porter, A. C. (1989). Do textbooks dictate the content of mathematics instruction in elementary schools? *American Educational Research Journal, 26*(3), 403-421.

Porter, A., Floden, R., Freeman, D., Schmidt, W., & Schwille, J. (1988). Content determinants in elementary mathematics. In D. Grouws & T. Cooney (Eds.), *Perspectives on research on effective mathematics teaching,* (pp. 96-113). Hillsdale, NJ: Erlbaum.

Schwille, J., Porter, A., Belli, G., Floden, R., Freeman, D., Knappen, L., Kuhs, T., & Schmidt, W. (1983). Teachers as policy brokers in the content of elementary school mathematics. In L. Shulman & G. Sykes (Eds.), *Handbook of teaching and policy,* (pp. 370-391). New York: Longman.

The teacher log developed by Michael Knapp and his colleagues (referenced in endnote 3) was heavily influenced by the content determinants research program.

6 Bill Connett at the Institute for Survey Research reports that a few such instruments are beginning to be developed and that there is some history of their design and issue, but that what we have in mind is more extensive than existing instruments. Still, our consultants at ISR believe that the log is a realistic idea and are excited to work on designing it and getting it to work well. This is, according to them, ground-breaking work in survey design.

7 In the classroom observations, researchers did not systematically check student attendance. Thus, for the lessons for which no students were marked absent, it is not clear whether teachers did not have any absent students or whether they just skipped that part of the log.

8 As we design the study, we need to find a consistent way to match student names across the teacher log and various other data collection sources at the student level, such as assessments. Since schools and teachers often use class lists that contain multiple small differences in how student names are recorded (e.g., various versions of nicknames, misspellings, changing surnames, etc.) maintaining matched records over a very large sample of students will present a host of challenges.

[9] The choices of grouping formats presented to teachers does not exhaust the variety of ways teachers typically organize students. For example, groups containing 10 or more students and pairs of students are not described well by the term "small group" or the term "whole group."

[10] In cases in which the classroom teacher is not teaching the lesson and is also not in the classroom, the log serves to capture data on the intended lesson and not the enacted lesson. This is also true in instances where the teacher reports to the best of her knowledge on the activity of other teachers working in her classroom.

[11] Our goal for this version of the teacher log was to learn how teachers themselves represented instructional topics. Teachers filled in an open ended topic field with their own descriptions of the lesson content. Teachers were not presented with any list of topic names to select from. In the next pilot we intend to present teachers with a list of topics from which they will select. Considerations for how to define the topic field are discussed in greater depth later in the report.

[12] In future analyses we will attempt to relate teacher and researcher responses, both in number of topics used per lesson (it seems researchers broke lessons down into more topics than did teachers) and in the words used to describe topics for the same lessons.

[13] We are considering using a collection of instruments as our "log" for the actual study. Thus, something like the way a teacher uses materials or implements a curriculum document could be part of a survey or a less frequent set of log questions, rather than included in the daily log.

[14] In addition, the use of the classroom teacher category was also intended to be a way to characterize the other teachers in the classroom. The researcher marked the teacher as the classroom teacher to differentiate her from the other teachers in the classroom who were designated as "other." Future versions of the log which better capture other adults in the room and what they are doing during the lesson may provide for a more narrow, more accurate definition of the classroom teacher and avoid future measurement error.

[15] Andrew Porter and his colleagues (see endnote 5) reported substantial researcher burden in their work implementing and maintaining regular log reports for a group of less than fifty teachers. One of the difficulties they confronted was keeping data collection up to date. Although the Web-based technology might help somewhat with the problem of "getting the data" since all entries are immediately entered into the data base, the research team would still need to contact teachers who fell behind in reporting. Porter and colleagues also found that they needed to review all log entries in order to identify teachers who were filling out the instrument incorrectly. Again, while the Web-based technology might facilitate the analysis of the data, since all data is automatically entered, it is still the case that teachers who exhibit confusion would need to be contacted by members of the research team.

APPENDIX

TEACHER LOG PILOT
HELP PAGES

Introduction to the Teacher Log

The teacher log is designed to create a daily record of your teaching in mathematics and reading. Our goal is to create a Web questionnaire that allows you to record certain features of your teaching in ways that yields accurate information about the content, materials, purposes, and specific activities that make up your lessons. We appreciate any feedback you can give us about the form or content of the log. This is still a work in progress and we are looking forward to your advice on how to ask clearer and more sensible questions. We also hope you will have ideas about how to make the log easier to use. Eventually, our plan is to have a computer log that takes no more than ten minutes to fill out each day. We expect that this version will take a bit longer since it is still rough and awkward in many places. It is also likely that the log will require some learning and after a number of uses will become much faster. Our aim is to take advantage of what we learn in this pilot work to develop a log that is quick, useful, and not overly burdensome to teachers.

This manual explains the major parts of the log and procedures for using the log. The same information is available as an online help page. Please note that the help pages for reading and mathematics are quite redundant, except for the TOPIC help section. These sections provide examples of mathematics and reading activities that illustrate what we are trying to learn about the tasks that students work on in each topic area.

Some General Guidelines

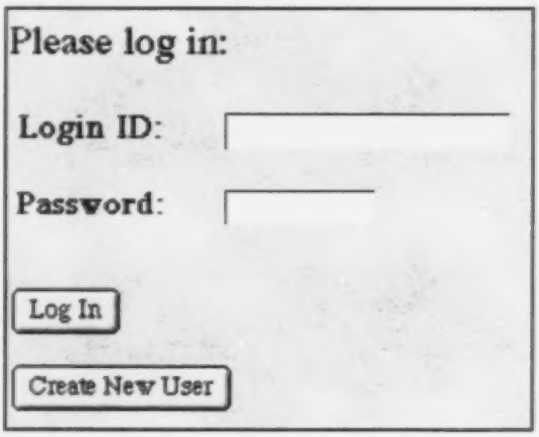
Because this is a pilot instrument, the teacher log is still fairly fragile. Some things don't work quite the way we want them to, and we will surely find some "BUGS" during this pilot. Please bear with us. If you run into a problem, please feel free to call or email with any questions you have.

A couple of warnings:

- Please DON'T use the BACK button to navigate through the questionnaire. It will cause you to lose data. Every page has navigation links. These usually include a NEXT link as well as links to specific pages you've already completed. Use these links to go back to previous work if you need to change something.
- SAVE actually sends your information off to the database. Until you save, your data is just sitting on your computer, so be sure to SAVE at every opportunity. There will be a SAVE button at the end of every sequence of data you record. The SAVE button is always in a RED bar.

SECTION 1: GETTING STARTED

The first screen you will see is the log-in, shown below. The very first time you use the log, you'll need to click on "CREATE NEW USER" and give yourself a name and password, as well as filling in information about your school.



Create New User

The screen below (Figure 1) is the first thing you'll see when you are getting started. For the LoginID and Password, choose things you will remember easily. We don't anticipate security problems - the name and password are primarily our way of hooking you up to your own class. So, your name and password don't need to be top secret. BUT DO REMEMBER THEM!

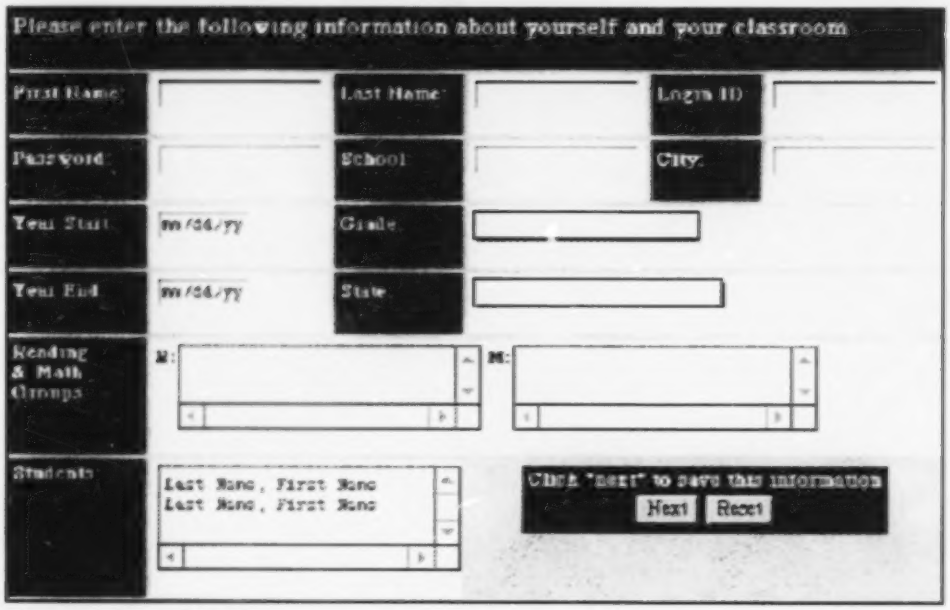


Figure 1: Create New User

BEST COPY AVAILABLE

For the pilot, please enter the dates 5/1/98 through 6/1/98 in the date boxes on this screen.

The reading and math groups are for you to specify groupings that you use regularly. Any name you give them is fine. The first box, with an "R" beside it, is for reading groups; the second with the "M" for math groups. If you don't use groups routinely, just skip these two boxes. Each group name should start on a separate line (use the "return" key on your computer to make a new line.)

Student names should go in as indicated, last name first, separated by a comma from the first name. Each student should be on a separate line. You can write over the "Last name, First name" lines already there, or erase them before you start. For purposes of the pilot, it is not important that we have the students' full or even real names. You can abbreviate or simplify as long as you know who each student is. (For example John Miller might be entered M, John.) On another screen, you will be able to designate which students are in which groups for both reading and mathematics.

Click on "NEXT" to move on.

Create Records

The next screen you see looks like the one below (Figure 2). Creating the records is a two step process in the software we are using. This screen presents step two. Don't go BACK to make corrections - you will get a chance to edit the records on future screens. All you can do here is click "Create Records." Once you click on this button, you have a database containing a list of your students. The next few screens will let you assign students to groups and make changes to your student or group lists. Go to the next screen.

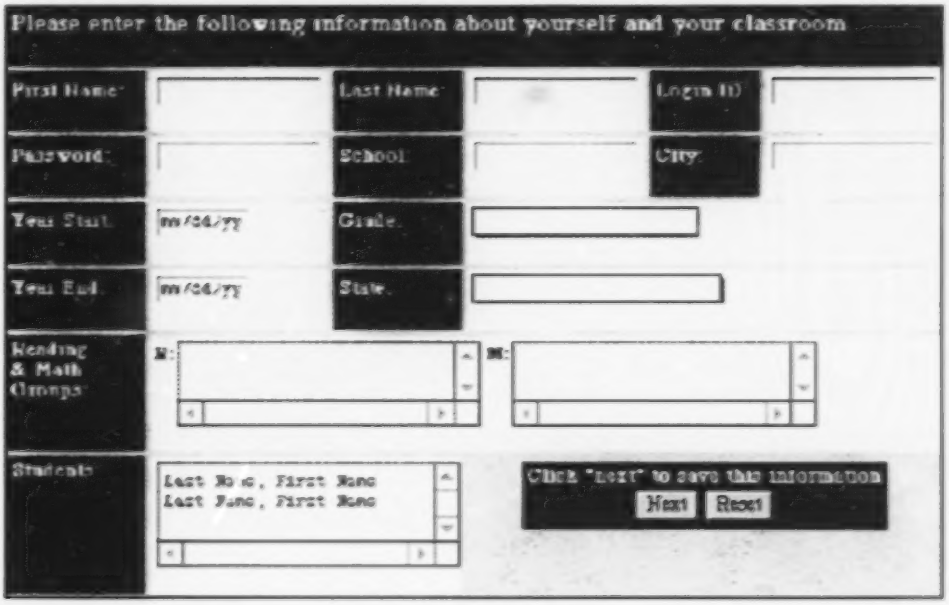


Figure 2: Create Records

BEST COPY AVAILABLE

Review and Edit

The screen below (Figure 3) shows what you will see after you create your records. If you need to make changes to the student name or the start and finish dates, pick "Edit Students." For later versions, the start and finish dates will be used when students start school late or leave early. That way, we'll know how long each student was actually in the classroom. For the pilot, you probably won't have any changes in your enrollment during this two week period, but if we do, this is how you report it.

Remember DON'T use the BACK button to navigate - you may lose data!

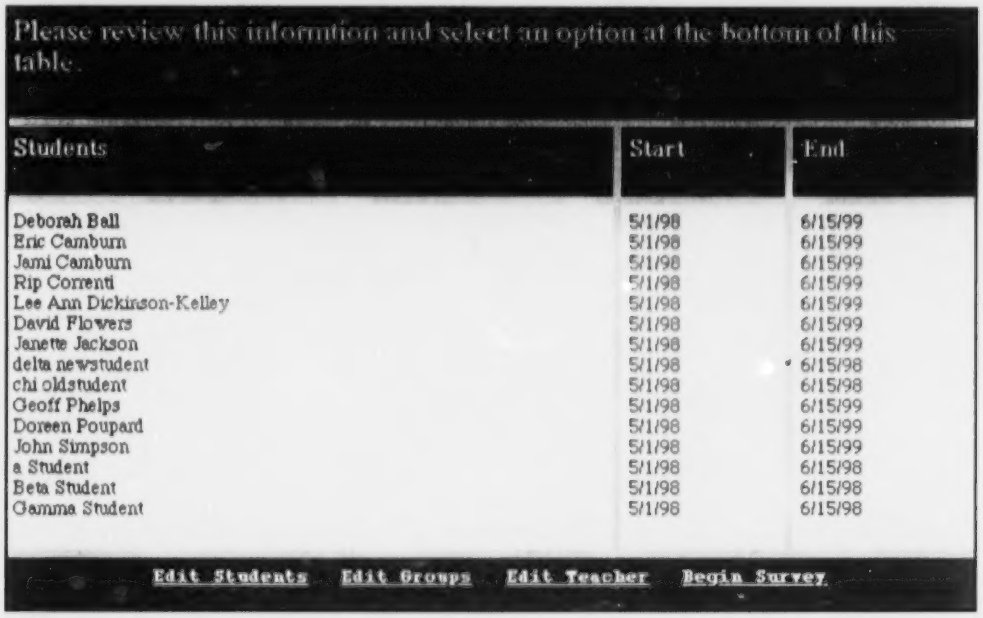


Figure 3: Review and Edit

Some of the editing features are not working correctly in the first version of the log. In particular, if you add students or groups after you first create your class (Figure 1), they will not show up on later screens. We are working on correctingthis problem . In the meantime, we will work with you to get all of your students entered correctly the first time through.

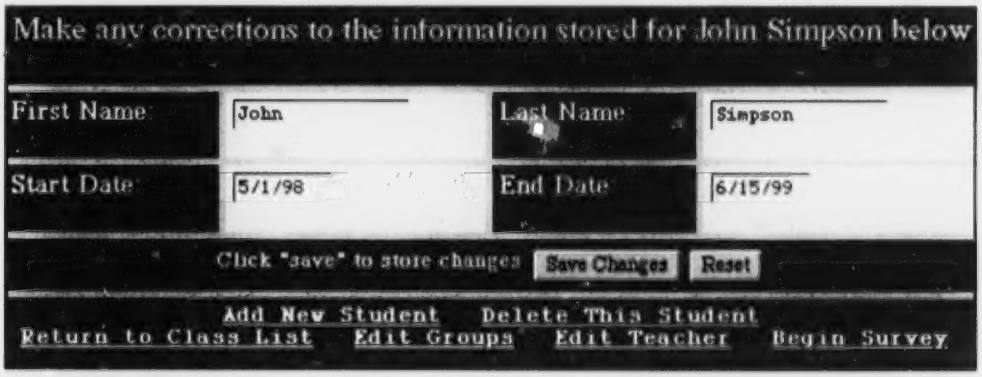


Figure 4: Edit Student

41 BEST COPY AVAILABLE

Edit Teacher

The screen for editing your teacher record is similar to the initial screen on which you first entered information to create a new user (Figure 1). You can check the information you entered there, and make changes as needed.

Edit Groups

To assign students to groups, select the "Edit Groups" option. When you see the screen below, select one of the groups to start assigning students.

Group (click on name to edit)	Subject	Members
Jacks	Math	none
Kings	Math	none
Queens	Math	none
Bluebirds	Reading	none
Cardinals	Reading	none
Warblers	Reading	none

Add New Group Edit Students Edit Teacher Begin Survey

Figure 5: Edit Groups

Define Groups

Here's what the screen looks like for defining groups (Figure 6.) The list of your students appears on the right, and you check (by clicking with the mouse) the students who belong in the group. Once you have made the assignments to this group, click on the "Return to Group List" button to select another group and to see what you have done.

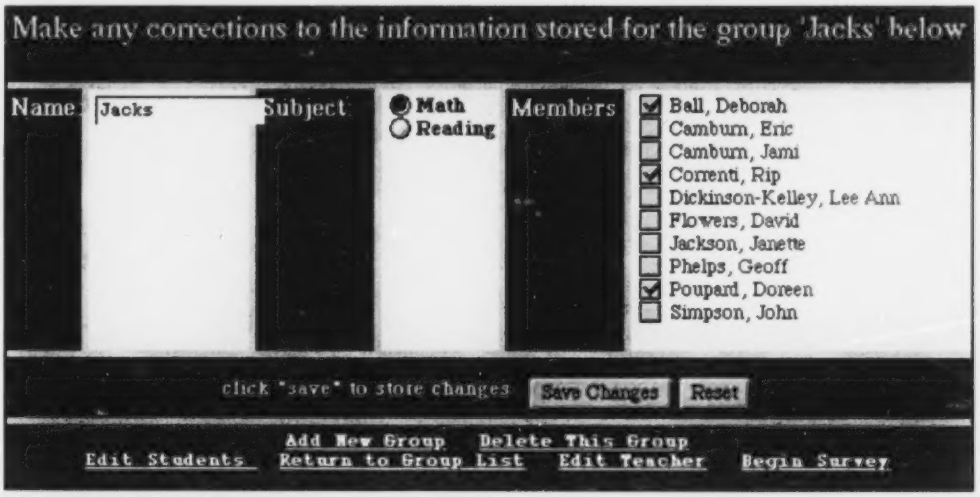


Figure 6: Define Groups

42 BEST COPY AVAILABLE

Review Groups

Below (Figure 7), you can see a complete group assignment for this class. Once you are satisfied with your groups, you can look at student records, or at your teacher record for editing. When you have finished group assignments or editing, click on the "Begin Survey" link to start reporting your lessons.

If for some reason you don't want to enter your lesson now (e.g., you just put in all your students and you're ready for a break), use the LOGOUT link (underlined and white in the bottom box) to leave the system.

Group (click on name to edit)	Subject	Members
Jacks	Math	Ball, Deborah Comenti, Rip Poupard, Doreen
Kings	Math	Camburn, Eric Dickinson-Kelley, Lee Ann Flowers, David Simpson, John
Queens	Math	Camburn, Jami Jackson, Janette Phelps, Geoff
Bluebirds	Reading	Jackson, Janette Phelps, Geoff Poupard, Doreen Simpson, John
Cardinals	Reading	Comenti, Rip Dickinson-Kelley, Lee Ann Flowers, David
Warblers	Reading	Ball, Deborah Camburn, Eric Camburn, Jami

Add New Group Edit Students Edit Teacher Begin Survey

Figure 7: Completed Groups

SECTION 2: DAILY ENTRY - READING AND MATHEMATICS

The first thing you need to do when you begin to fill in the log for a given day is to put in the date you're reporting, and the number of reading and mathematics lessons you taught. **FOR THE PILOT, at least for the first couple of times you use it, PLEASE REPORT ONLY ONE LESSON AT A TIME.** Even if you want to report both a mathematics and a reading lesson for a given day, please do them individually (complete one, log out, log back in, then do the other.) There are a couple of bugs in the system which have caused us to lose data when we put in several lessons at once. We also have found in using the log ourselves that for starters one lesson is about the right chunk of work. On subsequent uses we got much faster and were able to report on a days worth of reading and math lessons with much greater ease.

BEST COPY AVAILABLE

The screen for reporting the date and the number of lessons is shown in Figure 8. By "lesson," we are thinking of any discrete time during the day when you teach reading or mathematics. For example, you may use calendar at the start of the day as a mathematics or a reading lesson; or you may have sustained silent reading after lunch. Both of those would "count" as lessons, in addition to more specific curriculum driven lessons you teach. Since many days will have more than one mathematics or reading lesson, we suggest that on busy days you choose just one of these lessons to report on. At least for the first few passes at filling out the log this will substantially cut down on the burden of recording your instruction for the day.

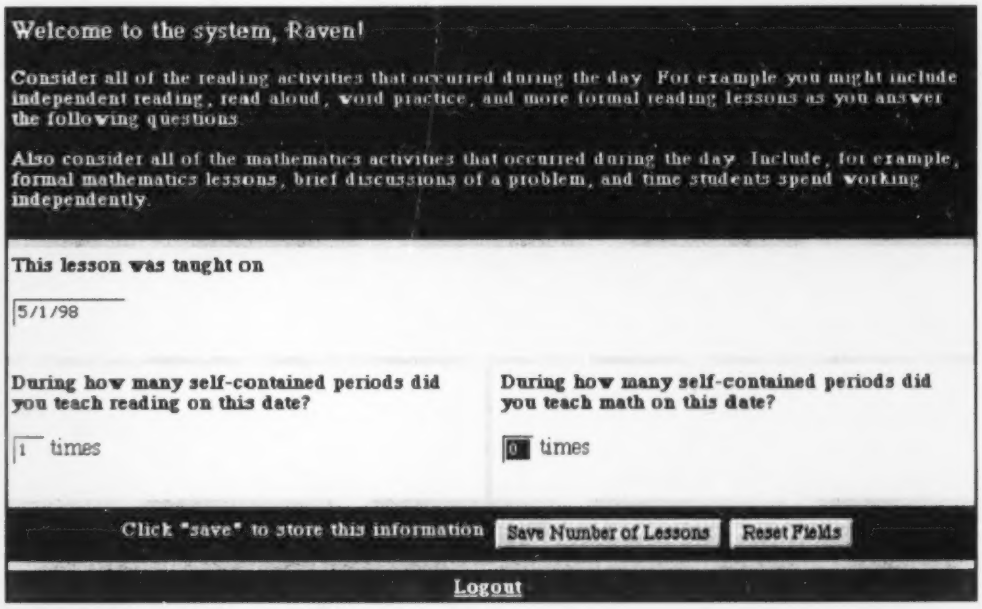

Welcome to the system, Raven!
Consider all of the reading activities that occurred during the day. For example you might include independent reading, read aloud, word practice, and more formal reading lessons as you answer the following questions.
Also consider all of the mathematics activities that occurred during the day. Include, for example, formal mathematics lessons, brief discussions of a problem, and time students spend working independently.
This lesson was taught on
5/1/98
During how many self-contained periods did you teach reading on this date?
1 times
During how many self-contained periods did you teach math on this date?
0 times
Click "save" to store this information
Save Number of Lessons
Reset Fields
Logout

Figure 8: Date and Number of Lessons

Reading Lesson Profile

The first screen in reporting a reading lesson provides a profile of how the students were grouped and what topics were taught.

Grouping: We are interested in learning how you worked with the students, whether as a whole class, in groups, or individually. You can pick more than one of these choices.

Topic: **TOPIC IS A MANDATORY FIELD!** If it is blank, the system gets very confused. In future versions we will have the system warn you of this, but for now, it's up to us to remember to get a topic entered.

Topic refers to the content of the lesson. Eventually, we want to provide a list of topics from which teachers can choose. As a first step, we are interested in learning how you think about topic and what words you use to describe the content of your lessons. In our development work on the log, we have tried to create comprehensive lists of topics for both mathematics and reading. We have found that

BEST COPY AVAILABLE

identifying and naming topics is very difficult! Often for any given lesson a number of topics are covered at the same time. Sometimes it is not clear what words best describe a topic. We have no set idea of the best way to characterize topic coverage. At this point in the work your best attempt to describe the content of the lesson will be very helpful. For each topic you identify, the log will ask for information about the purpose, materials, and activities associated with teaching that topic. **The maximum number of topics** for a single lesson (for the pilot) is five. Is this enough? Too many? There are comment boxes in the log that you can use to make suggestions about topics and other issues that come up.

Separate topics are entered on separate lines in the box provided. Create a separate line using the return key, just as in a word processing program.

DON'T CLICK DONE YET! The next four screens - Teacher, Students, Comments, and Done - are actually on the same Web page as this topic screen. That means you can either use the white and red navigation links (named Lesson, Teacher, Students, Comments, Done) or scroll down with the scroll bar. It also means you can go back and forth on this page until you get it the way you want it. THE DONE BUTTON IS TO SAVE ALL OF THIS WORK when you've finished with everything up through the Comment screen. Go on to Teacher next, and work your way through this set of links sequentially until you reach Done.

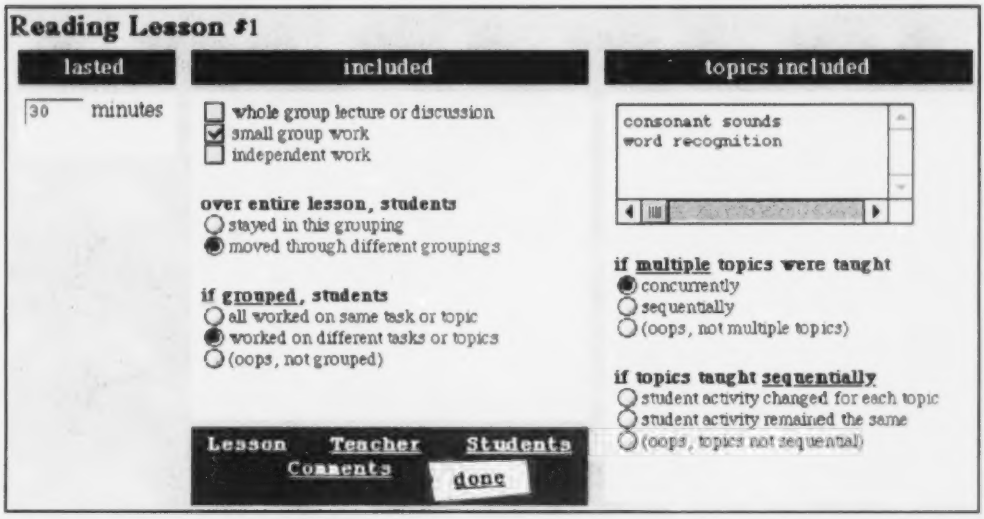


Figure 9: Reading Lesson Profile

Teachers

This form (Figure 10) is fairly straightforward. You can indicate that more than one teacher - or other adult - was in the classroom. We haven't figured out quite how to specify who was doing what as they worked in the classroom, so please use the Comments form below to clarify anything that seems poorly described. Don't click DONE yet! Go on to Students next.

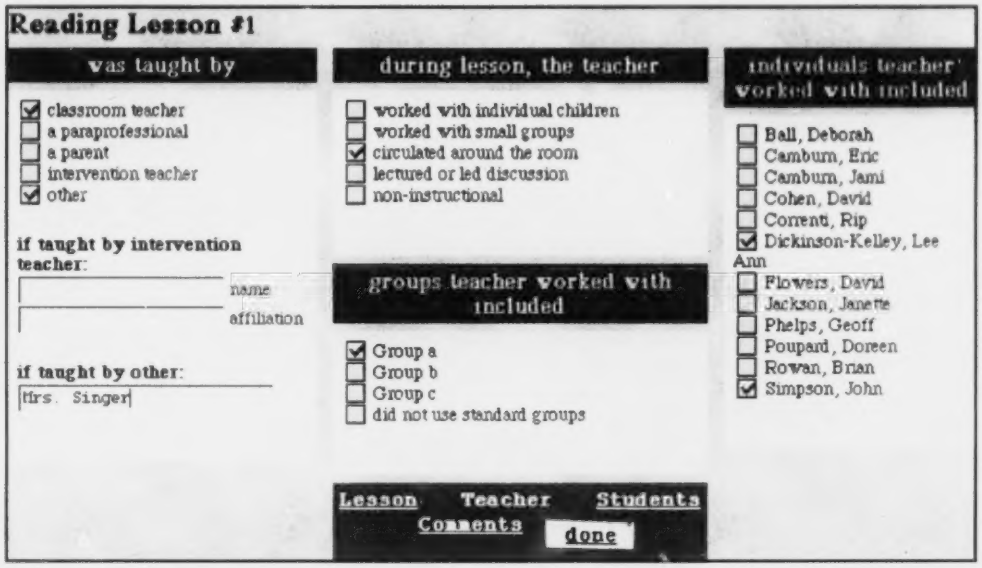


Figure 10: Teacher Activity

Students

This is also a fairly straightforward screen (Figure 11.) Who was absent? (The maximum number of absences you can enter for a single lesson during the pilot is five. Can you think of many times when you had more then five students absent for a lesson?) What kind of activity did students engage in?

Checking one of the boxes from the list of student activities is mandatory. Leaving them all blank causes the system to do odd things, and may result in losing data. Again, we will have a system warning to this effect in later versions, but for now, it's up to you to remember, and us to remind you, to fill in at least one student activity.

More questions about the absent students, and about the activity will follow. Don't click DONE yet! Go on to Comments next. If you want to go back to previous screens, you can scroll back through the last three screens, but don't use the BACK button - you'll lose your data.

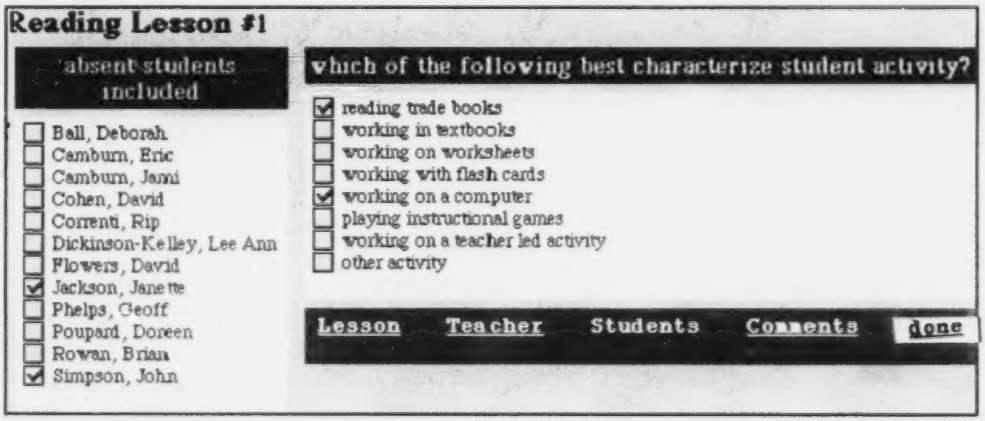


Figure 11: Student Activity

BEST COPY AVAILABLE

Comments

Use this comments box (Figure 12) to clarify any of the information provided above. You may want to write a short summary of how the lesson was structured, or note the role of adults who were helping in the classroom. You can also use this space to help us understand how we need to change or expand the choices you were given above. You can scroll back up the page to review what you have done. None of this data is saved until you go to DONE and click "Save Lesson Info" as shown below (Figure 12). Until you save, you can scroll back up the page and make changes, but, at the risk of being redundant, don't use the back button.

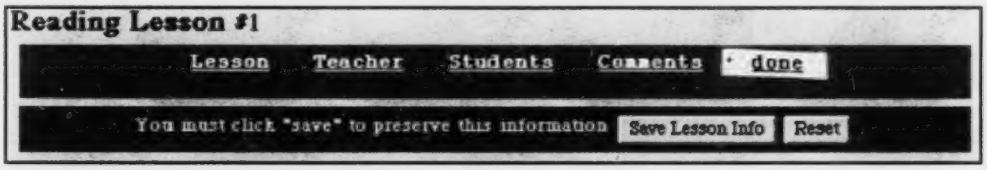


Figure 12: Saving Lesson Data

Absences

After your basic lesson data is entered, you will see a screen to record why students missed the lesson (Figure 13). When you have complete this form, click on "Save Absence Info" to save the data.

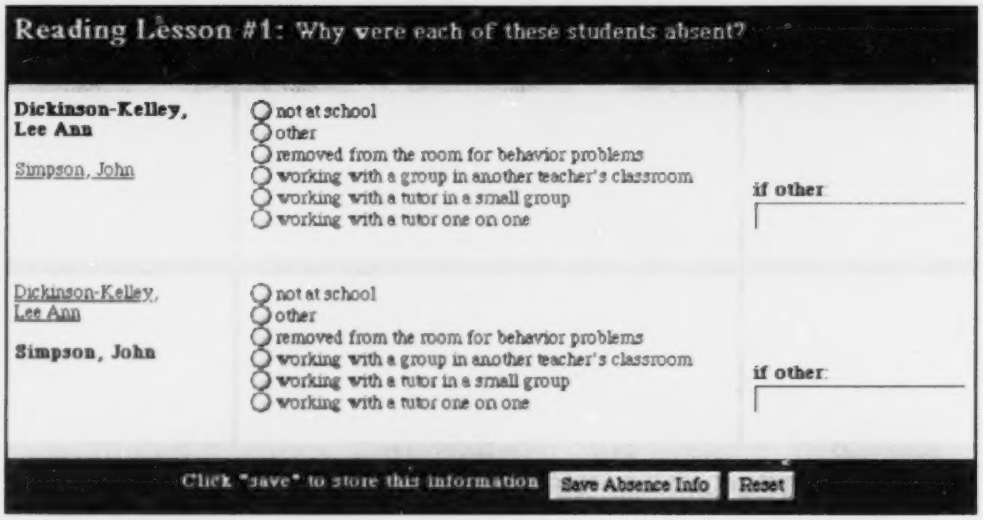

Reading Lesson #1: Why were each of these students absent?
Dickinson-Kelley, Lee Ann
Simpson, John
not at school
other
removed from the room for behavior problems
working with a group in another teacher's classroom
working with a tutor in a small group
working with a tutor one on one
if other:
Dickinson-Kelley, Lee Ann
Simpson, John
not at school
other
removed from the room for behavior problems
working with a group in another teacher's classroom
working with a tutor in a small group
working with a tutor one on one
if other:
Click "save" to store this information
Save Absence Info
Reset

Figure 13: Absences

BEST COPY AVAILABLE

Topics - Student Work

Our goal for this section is to learn more about the kind of work students were doing. We have tried to think of categories that will work across many different kinds of activities and topics. For many reading topics, you may have more than one goal in mind as you teach. The categories provided below (Figure 14) look at reading topics from the point of view of what students were doing during the lesson. For example, when reading a book, students might primarily be practicing word recognition one day, and on another day they might be answering questions that require them to draw on their background knowledge. For the lesson you taught, which of categories below best characterizes the activity of your students. The numbers allow you to rank order the importance of each category (1 is most important, 4 is least important). It is ok to have two categories ranked the same, and to designate some as not applicable. Each topic has a screen like this. DON'T click DONE until you have finished all of your topics. These are all on the same Web page, so you can scroll up and down, or use the white underlined word links to navigate among the topics.

We are very interested in your feedback about the usefulness and completeness of these categories. One of the big problems in studying teaching is that the most obvious things about instruction such as groupings and activities often say little about the nature of the content that students are working on. For example worksheets can focus on letter and sound recognition or they can focus on comprehension. We are very interested in learning about the nature of student work and this set of questions is our attempt to get a rough sense of the nature of the content in the lesson. We are eager to figure out ways to make this work, but given the complexity of many lessons we are also aware that this may not work the way we hope. If you have taught a lesson or a topic within a lesson for which these categories make no sense, please let us know how we could make these categories more useful, or how we could change what we are asking to better reflect what you are doing. A more complete description of how we are thinking about these particular categories for reading topics is given in the table below.

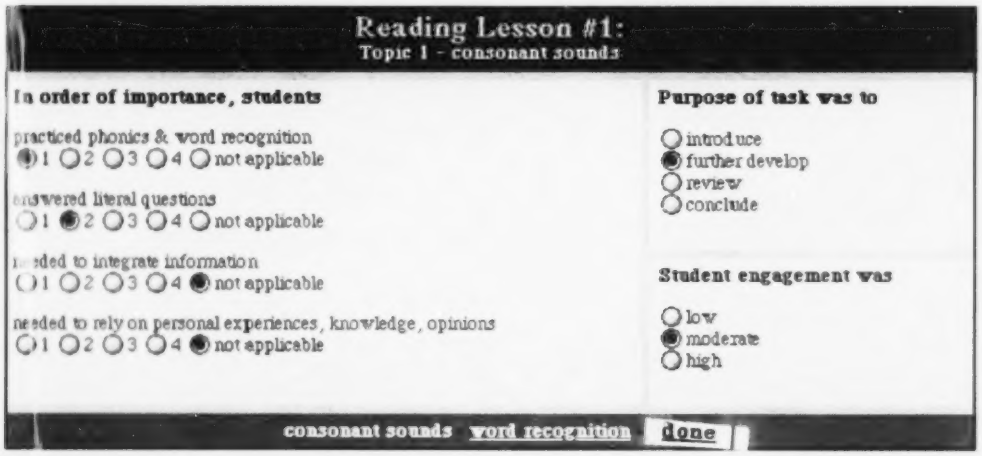


Figure 14: Topics

Table of Examples for Reading Topic Analysis

Practiced phonics and word recognition

Students practiced phonics and word recognition. Phonetic work requires students to listen for, produce, or read letters and combinations of letters and sounds. Word recognition requires students to practice common words, decode less familiar words, or read from text.

Examples of this kind of student activity

1. Circle all of the words that have a short /à/ sound like you hear in cat. bad - mud - lane - at - car.
2. Students are directed to read each word and circle the word that has the sound they hear at the beginning of "around": ache, agree, any.
3. The teacher asks a students to point to the words that begin with /n/ sound: nurse, lid, crate, nest.
4. The teacher asks students to think of a word that begins with /b/ and ends with /t/ and write down their answers.
5. Students read from books or other text without answering questions or engaging in conversations about what they have read.

Answered literal questions

Students answered literal questions about characters, setting, or other features of a story or poem. Students did not need to interpret text, make inferences, or draw on their own background experience to answer these questions.

Examples of this kind of student activity

Students are asked to read a story and answer questions.

Since November, Makiko Ogawa (mu-ke-ko o-ga-wu) had gone ice-skating every Saturday at the indoor ice skating arena. Makiko always went skating with her older friend Robert and her little sister, Lin. Robert was a good skater and today he had promised to teach the two girls how to skate backwards. Makiko wished that for once she could leave Lin at home. She really wanted to learn and how to skate backwards. She was worried that Robert would spend the entire morning helping Lin.

i. What kind of skater is Robert?
ii. What did Makiko want to learn?
iii. Who always went skating with Makiko?
iv. What was Makiko worried about?

Needed to integrate information

Students needed to integrate information to answer questions about characters, setting, or other features of a story or poem. Students needed to "read between the lines," perhaps inferring or summarizing information in order to answer questions.

Examples of this kind of student activity

Students are asked to read a story and answer questions.

Since November, Makiko Ogawa (mu-ke-ko o-ga-wu) had gone ice-skating every Saturday at the indoor ice skating arena. Makiko always went skating with her older friend Robert and her little sister, Lin. Robert was a good skater and today he had promised to teach the two girls how to skate backwards. Makiko wished that for once she could leave Lin at home. She really wanted to learn and how to skate backwards. She was worried that Robert would spend the entire morning helping Lin.

i. What is an arena? Look the word up in the dictionary to see if you are correct.
ii. Who is the oldest person in the story? Who is the youngest?
iii. Why did Makiko want to leave Lin at home?

Needed to rely on personal experiences, knowledge, opinions

Student could not complete work by just examining the text. Students needed to rely on their own personal experiences, knowledge, or opinions in order to answer questions and complete tasks.

Examples of this kind of student activity

Students are asked to read a story and answer questions.

Since November, Makiko Ogawa (mu-ke-ko o-ga-wu) had gone ice-skating every Saturday at the indoor ice skating arena. Makiko always went skating with her older friend Robert and her little sister, Lin. Robert was a good skater and today he had promised to teach the two girls how to skate backwards. Makiko wished that for once she could leave Lin at home. She really wanted to learn how to skate backwards. She was worried that Robert would spend the entire morning helping Lin.

i. Why do you think Robert took the two girls skating every Saturday?
ii. What would you change if you decided to write this story from Lin's perspective?
iii. What does it feel like to ice-skate backwards?
iv. What is your favorite sports activity?
v. Do you have a special older friend like Robert?

Once you have entered this topic information, you need to save your work before moving on to materials. **BE SURE TO CLICK ON SAVE TOPIC INFO HERE** (Figure 15) to preserve your data! Until you click on this, your topic data has not been sent to the database. Next you will go on to the next topic you specified and go through the same screens for that topic, and for each additional topic you named. The underlined words in white in the middle bar are the topics designated for this lesson.

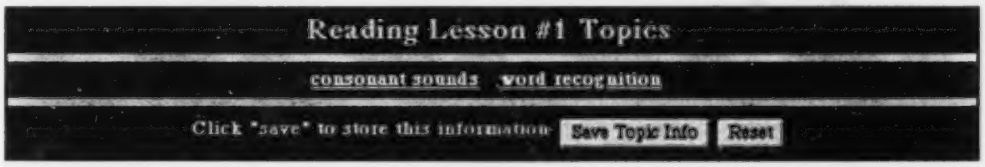


Figure 15: Save Topic Info

Materials

Here, we are interested in learning what materials you used. You will get different screens here depending on which items you chose under student activity. In the future, we will have the system create a list from what you input here (e.g., when you designate a text book) so that when you log in again, you can choose from that list. Be sure to click on Save Activity Info when you complete this screen. **Don't use the BACK button here, or this information will not be saved.**

If you entered more than one reading lesson, the log will move on to that lesson, going back to the screen shown in Figure 9 above. If you have no more reading lessons, but you do have mathematics lessons, the log will move into mathematics lessons next. If you have no more lessons, the log will go to the Review and Log Out screen.

BEST COPY AVAILABLE

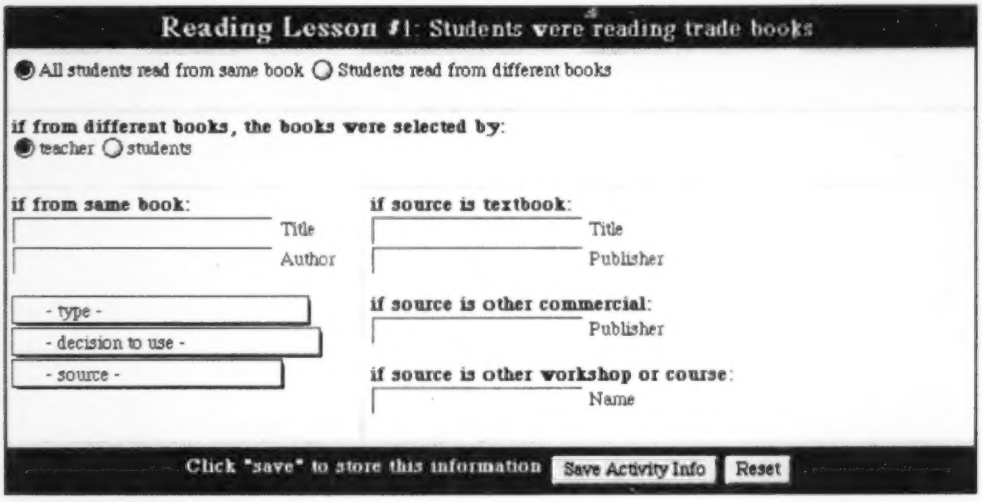


Figure 16: Reading Materials

Mathematics Lesson Profile

You can get to the screens for entering information about your mathematics lessons either after you complete a reading lesson, or directly from the initial screen on which you enter the number of lessons (if you enter no reading lessons). For purposes of the pilot, at least for the first couple of times you use it, please enter one lesson at a time. That way, you will enter your math lesson from the initial screen shown in Figure 8 above by entering 0 for the number of reading lessons, and 1 for the number of math lessons.

The first screen in reporting a mathematics lesson provides a profile of how the students were grouped and what topics were taught.

Grouping: We are interested in learning how you worked with the students, whether as a whole class, in groups, or individually. You can pick more than one of these choices.

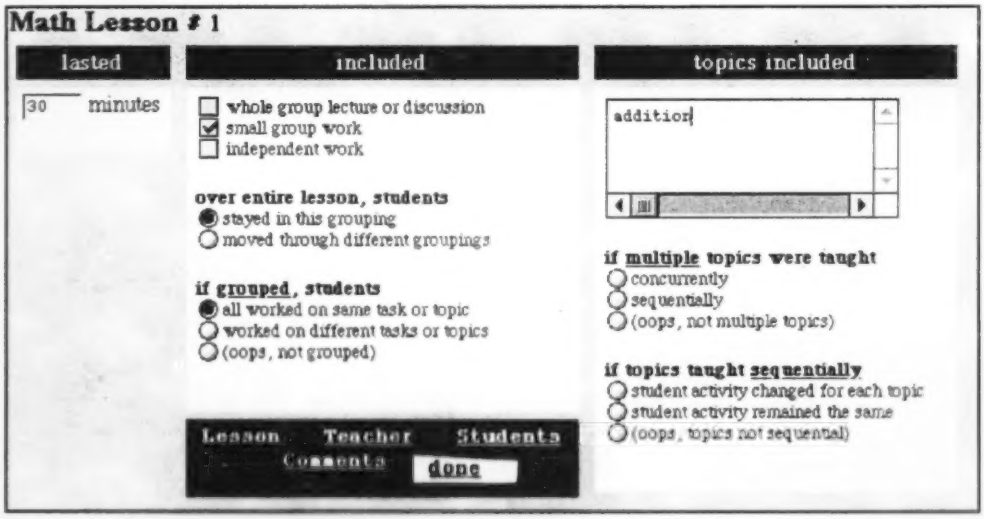


Figure 17: Mathematics Lesson Profile

BEST COPY AVAILABLE

Topic: **TOPIC IS A MANDATORY FIELD!** If it is blank, the system gets very confused. In future versions we will have the system warn you of this, but for now, it's up to us to remember to get a topic entered.

Topic refers to the content of the lesson. Eventually, we want to provide a list of topics from which teachers can choose. As a first step, we are interested in learning how you think about topic and what words you use to describe the content of your lessons. As we have worked to develop the log we have tried to create comprehensive lists of topics for both mathematics and reading, and we realize that this is not easy! Often at any given moment a number of topics are covered at the same time. Sometimes it is not clear what words best describe a topic. We have no set idea of the best way to characterize topic coverage. At this point in the work your best attempt to describe the content of the lesson will be very helpful. For each topic you identify, the log will ask for information about the purpose, materials, and activities associated with teaching that topic. **The maximum number of topics** for a single lesson (for the pilot) is five. Is this enough? Too many?

Separate topics are entered on separate lines in the box provided. Create a separate line using the return key, just as in a word processing program.

DON'T CLICK DONE YET! The next four screens - Teacher, Students, Comments, and Done - are actually on the same Web page as this topic screen. That means you can either use the white and red navigation links (named Lesson, Teacher, Students, Comments, Done) or scroll down with the scroll bar. It also means you can go back and forth on this page until you get it the way you want it. THE DONE BUTTON IS TO SAVE ALL OF THIS WORK when you've finished with everything up through the Comment screen. Go on to Teacher next, and work your way through this set of links sequentially until you reach Done.

Teacher

This form is fairly straightforward. You can indicate that more than one teacher - or other adult - was in the classroom. We haven't figured out quite how to specify who was doing what, so please use the Comments form below to clarify anything that seems poorly described. On this form, you must select one of the items under "materials teacher used..." This is another one of those **mandatory fields** that you can't tell is mandatory, so be sure to fill it in. Don't click DONE yet! Go on to Students next.

Students

This is also a fairly straightforward screen (Figure 19). Who was absent? (The **maximum number** of absences you can enter for a single lesson during the pilot is five. Can you think of many times when you had more then five students absent for a lesson?) What kind of activity did students engage in?

Checking one of the boxes from the list of student activities is mandatory: leaving them all blank causes the system to do odd things, and may result in losing data. Again, we will have a system warning to this effect in later versions, but for now, it's up to you to remember, and us to remind you, to fill in at least one student activity.

More questions about the absent students, and about the activity will follow. Don't click DONE yet! Go on to Comments next. If you want to go back to previous screens, you can scroll back through the last three screens, but don't use the BACK button - you'll lose your data.

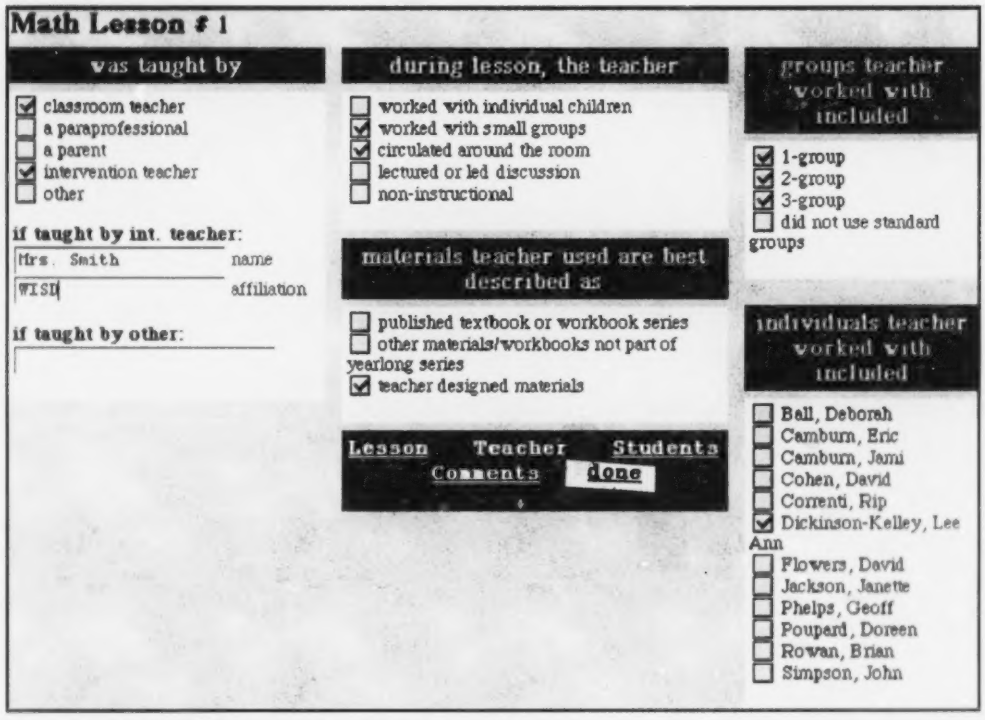


Figure 18: Mathematics Lesson Teacher Information

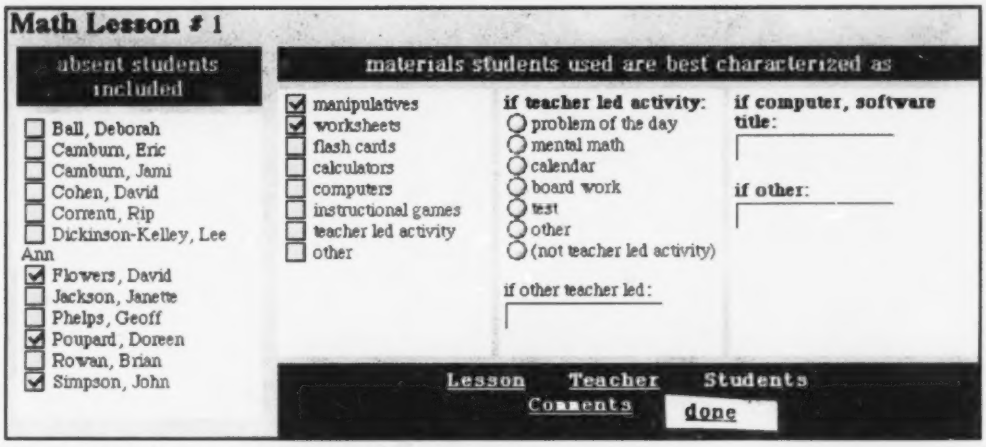


Figure 19: Mathematics Lesson Student Information

Comments

Use this comments box (Figure 20) to clarify any of the information provided above. You may want to write a short summary of how the lesson was structured, or note the role of adults who were helping in the classroom. You can also use this space to help us understand how we need to change or expand the choices you were given above. You can scroll back up the page to review what you have done. None of this data is saved until you go to DONE and click "Save Lesson Info" as shown below (Figure 21).

BEST COPY AVAILABLE

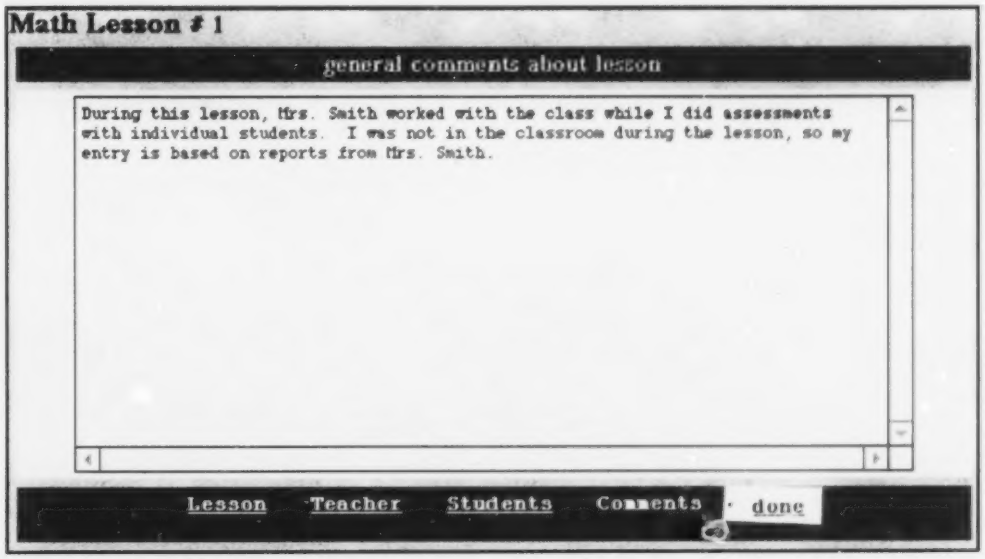

Math Lesson # 1
general comments about lesson
During this lesson, Mrs. Smith worked with the class while I did assessments with individual students. I was not in the classroom during the lesson, so my entry is based on reports from Mrs. Smith.
Lesson
Teacher
Students
Comments
done

Figure 20: Mathematics Lesson Comments

Done

Be sure to click on "Save Lesson Info" (Figure 21) to send all you've done on this lesson to the database.

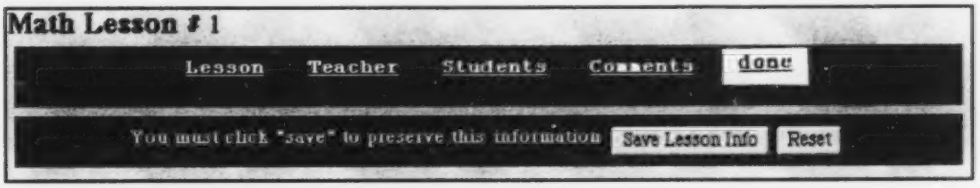


Figure 21: Save Lesson Info

Absences

After your basic lesson data is entered, you will see a screen to record why students missed the lesson. When you have complete this form, click on "Save Absence Info" to save the data (Figure 22).

Topics - Student Work

Our goal for this section is to learn more about the kind of work students were doing. We have tried to think of categories that will work across many different kinds of activities and topics. For many mathematics topics, you may have more than one goal in mind as you teach. The categories provided below (Figure 23) look at mathematics topics from the point of view of what students were doing during the lesson. For example, when practicing addition students might primarily memorize basic facts one day, and on another day they might practice computational procedures. For the lesson you taught, which of categories below best characterizes the activity of your students. The numbers allow you to

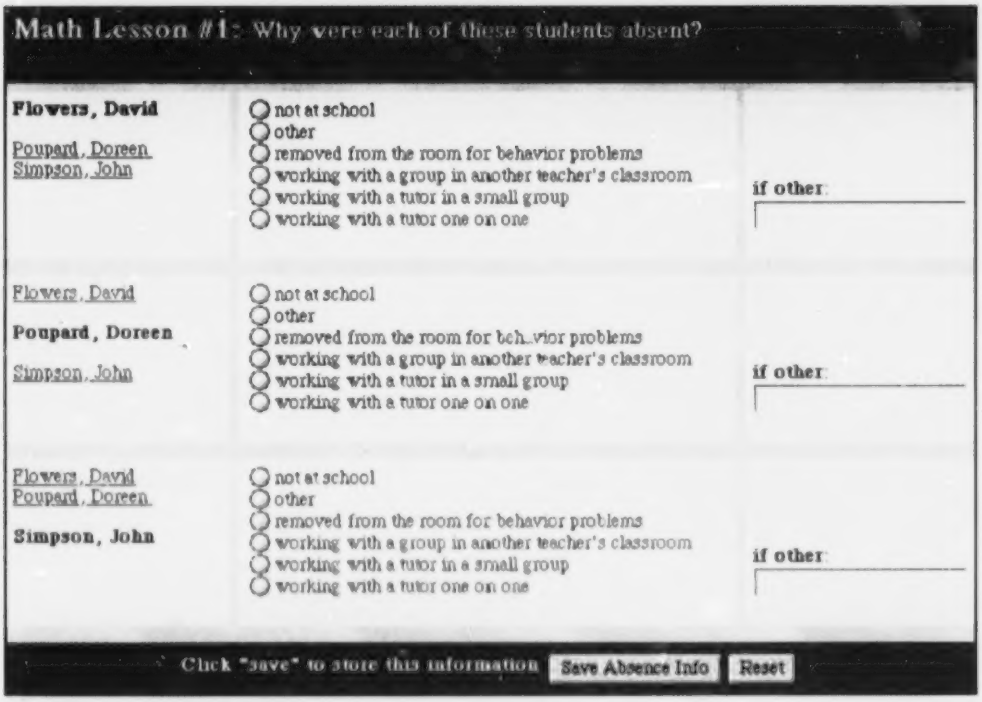

Math Lesson #1: Why were each of these students absent?
Flowers, David
Poupard, Doreen
Simpson, John
not at school
other
removed from the room for behavior problems
working with a group in another teacher's classroom
working with a tutor in a small group
working with a tutor one on one
if other:
Flowers, David
Poupard, Doreen
Simpson, John
not at school
other
removed from the room for beh..vior problems
working with a group in another teacher's classroom
working with a tutor in a small group
working with a tutor one on one
if other:
Flowers, David
Poupard, Doreen
Simpson, John
not at school
other
removed from the room for behavior problems
working with a group in another teacher's classroom
working with a tutor in a small group
working with a tutor one on one
if other:
Click "save" to store this information
Save Absence Info
Reset

Figure 22: Mathematics Lesson Absences

rank order the importance of each category (1 is most important, 4 is least important). It is ok to have two categories ranked the same, and to designate some as not applicable. Each topic has a screen like this. DON'T click DONE until you have finished all of your topics. These are all on the same Web page, so you can scroll up and down, or use the white underlined word links to navigate among the topics.

We are very interested in your feedback about the usefulness and completeness of these categories. One of the big problems in studying teaching is that the most obvious things about instruction such as groupings and activities often say little about the nature of the content that students are working on. For example worksheets can focus on memorizing basic facts or they can focus on computational procedures. We are very interested in learning about the nature of student work and this set of questions is our attempt to get a rough sense of the nature of the content in the lesson. We are eager to figure out ways to make this work, but given the complexity of many lessons we are also aware that this may not work the way we hope. If you have taught a lesson or a topic within a lesson for which these categories make no sense, please let us know how we could make these categories more useful, or how we could change what we are asking to better reflect what you are doing. A more complete description of how we are thinking about these particular categories for mathematics topics is given in the table below.

BEST COPY AVAILABLE

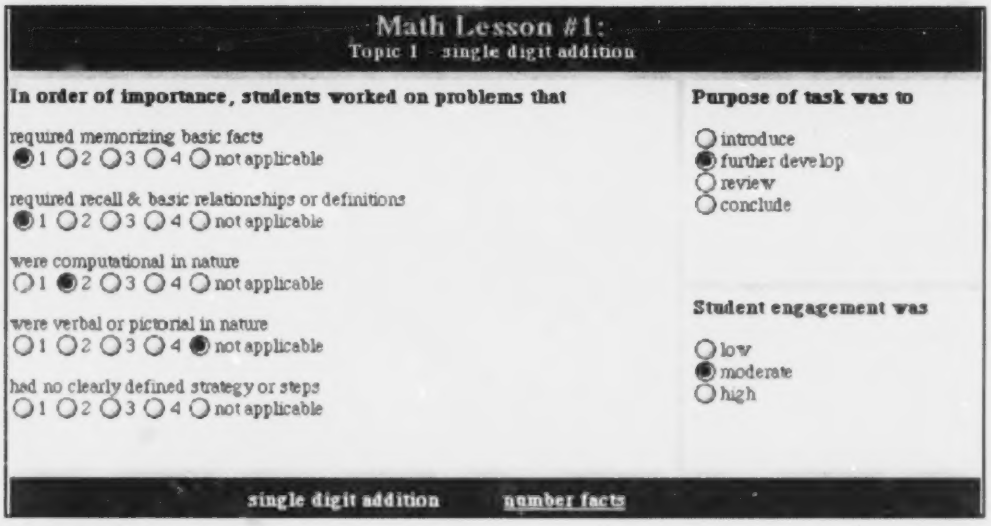

Math Lesson #1:
Topic 1 - single digit addition
In order of importance, students worked on problems that
required memorizing basic facts
1 2 3 4 not applicable
required recall & basic relationships or definitions
1 2 3 4 not applicable
were computational in nature
1 2 3 4 not applicable
were verbal or pictorial in nature
1 2 3 4 not applicable
had no clearly defined strategy or steps
1 2 3 4 not applicable
Purpose of task was to
introduce
further develop
review
conclude
Student engagement was
low
moderate
high
single digit addition
number facts

Figure 23: Mathematics Topic

Table of Examples for Mathematics Topic Analysis

Required memorizing basic facts

Students worked on memorizing basic facts. The work required students to supply an answer directly from memory and did not demand a series of computational steps.

Examples of this kind of problem

5 + 3 =

7 x 8 =

124 / 12 =

1/2 + 1/2 =

Required recall and basic relationships or definitions

Students worked on problems that primarily required recall and application of basic mathematical relationships or definitions.

Examples of this kind of problem

1. Which of the following numbers is even?

   4 56 12 25 32

2. Which of the following triangles is isosceles?

3. If a, b, c are real numbers and a > b, then ac > bc. True or False

4. Find the greatest common multiple of 36 and 45.

Were computational in nature

Students worked on solving computational problems. In order to solve these problems students learned or applied a standard series of computational steps.

Examples of this kind of problem

23 + 45 =

891 - 403 =

1/3 x 3 =

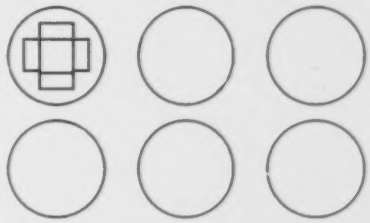
BEST COPY AVAILABLE

Were verbal or pictorial in nature

Students worked on word or pictorial problems. As a first step students may have rewritten these problems as computational problems. Students then learned or applied a standard series of computational steps.

Examples of this kind of problem

1. I had three dozen cookies and my mom took one dozen to work. How many dozen do I have left?
2. There are 30 children in our classroom. Two thirds of them are girls. How many are boys?
3. The class is planning a party. Only four placements will fit on each table. How many placements do they need?

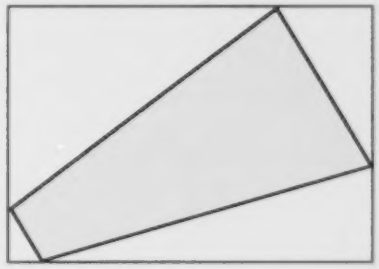


Had no clearly defined strategy or steps

Students worked on problems for which there was not a clearly defined strategy or set of computational steps. These problems can not easily be rewritten as computational problems. Different students may well have come up with quite different ways of working on the problems.

Examples of this kind of problem

1. I have pennies, nickels, and dimes in my pocket. If I take out three coins, how much money could I have? Find all the possible combinations.
2. Think of as many ways as you can for determining the area of the shaded region.



Next, you can move on to the next topic in your list of mathematics topics for this lesson, shown as a white link in the middle box below. If you have more than one topic, you can link to the next topic from the current topic screen, or you can scroll down the page. (All topics are on the same Web page.) **BE SURE TO CLICK ON SAVE TOPIC INFO on the next screen (Figure 24) to preserve your data!** Until you click on this, your topic data has not been sent to the database. Once you have saved the topic information, you will go on to the review and logout screens.

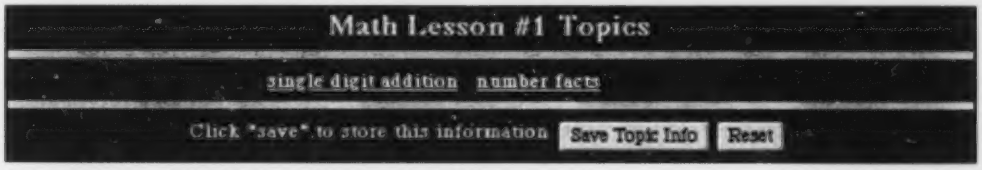


Figure 24: Save Topic Info

BEST COPY AVAILABLE

Materials

There are several options which you may see to describe the mathematics materials you used, depending on what you designated on the "Student" link on a previous screen. The possible choices are shown here (Figure 25)—you will get a slightly different screen for each item you chose. In the future, we will have the system create a list from what you input here (e.g., when you designate a text book) so that when you log in again, you can choose from that list. Be sure to click on Save Activity Info when you complete this screen. **Don't use the BACK button here, or this information will not be saved.**

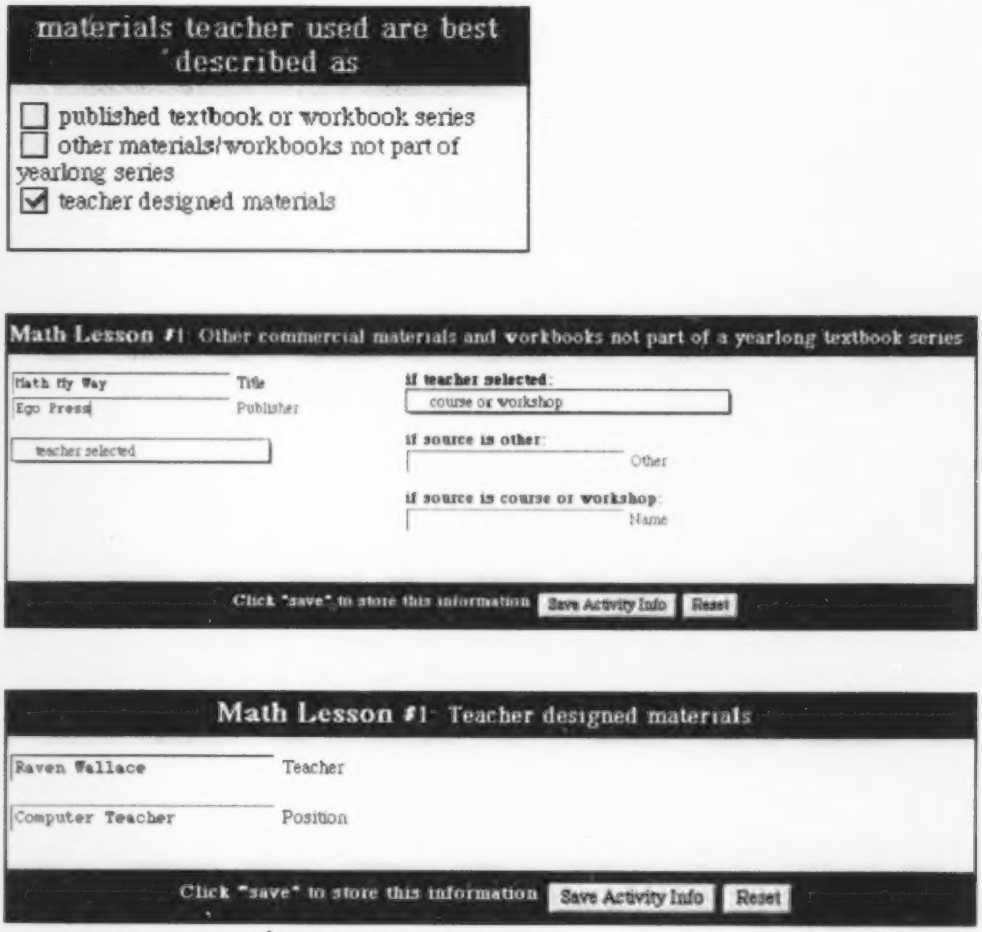


Figure 25: Mathematics Materials

BEST COPY AVAILABLE

REVIEW AND LOG OUT

When you get to the end of inputting lessons, you will have an opportunity to review your work and make changes. As you see from the sample screen, you can either look at a summary of the lesson, or log out (Figure 26).

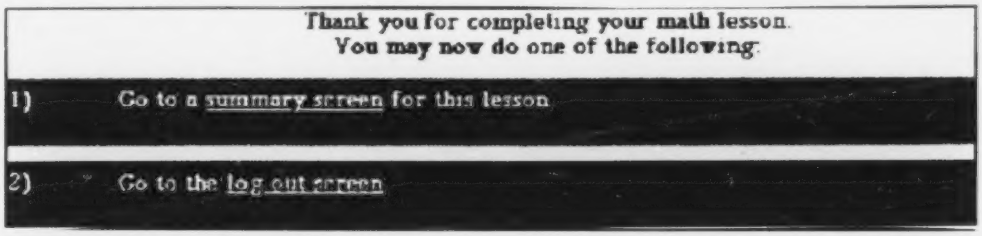

Thank you for completing your math lesson.
You may now do one of the following:
1) Go to a summary screen for this lesson
2) Go to the log out screen

Figure 26: Review and Log Out

If you choose to review your work, you will see a series of screens, starting with the screen below in Figure 27, which will allow you to make changes to what you have done. Eventually we want the review screen to be a true review which summarizes the days lessons on one screen. For now the review section repeats the same screens you have already seen allowing you to make changes. This is not ideal, since it does not allow you get a true overview. This is another area we intend to change in subsequent versions of the log. At any time during your review, even if you make no changes, you can click on the DONE button to save any changes you have made or to log out.

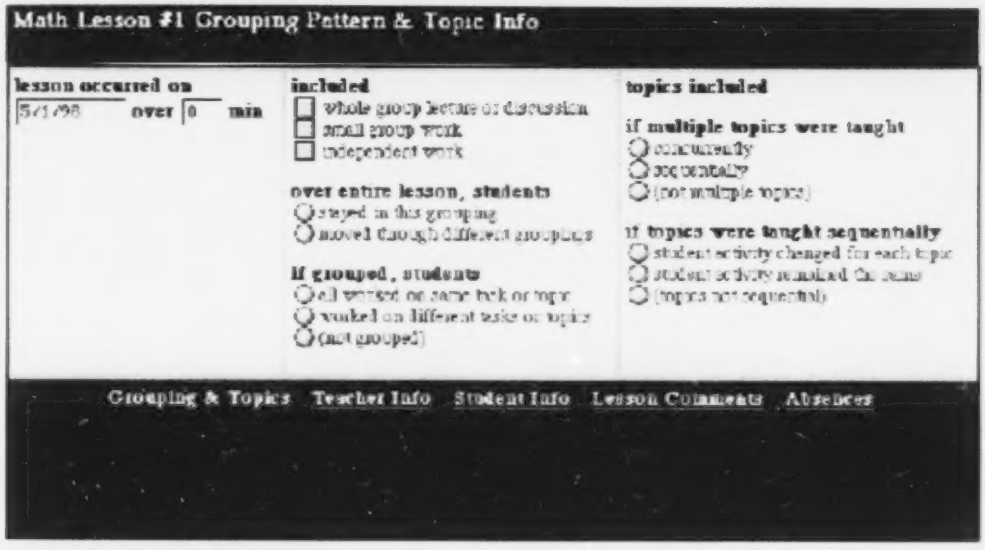


Figure 27: Review Math Lesson

BEST COPY AVAILABLE

When you click on DONE, you get the screen shown in Figure 28. "Save Corrections" will save any changes you have made and take you to the log out screen. You should also use "Save Corrections" if you have made NO CHANGES and you are ready to log out. Use RESET to clear any changes you made. You'll then need to use "Save Corrections" to get to the Log Out screen.

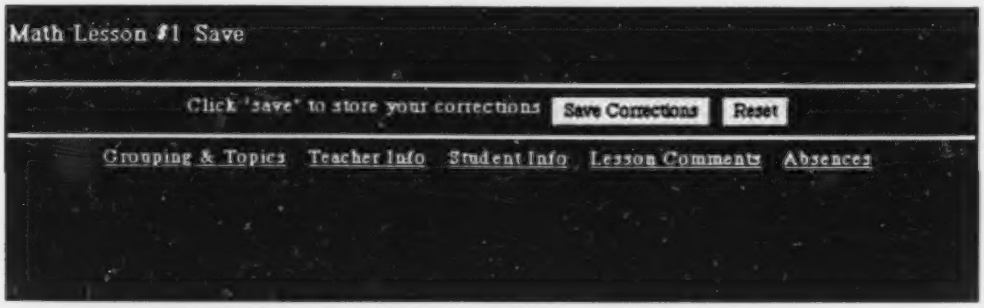


Figure 28: Return to Log Out

Finally, you can log out from the screen shown in Figure 29. This will take you back to the Log In screen. If you want to do another lesson, you can log back in immediately, or just quit Netscape and you're done! From this screen, we have been having some problems with making corrections to individual lessons, so avoid using that feature for now (number 1 below).

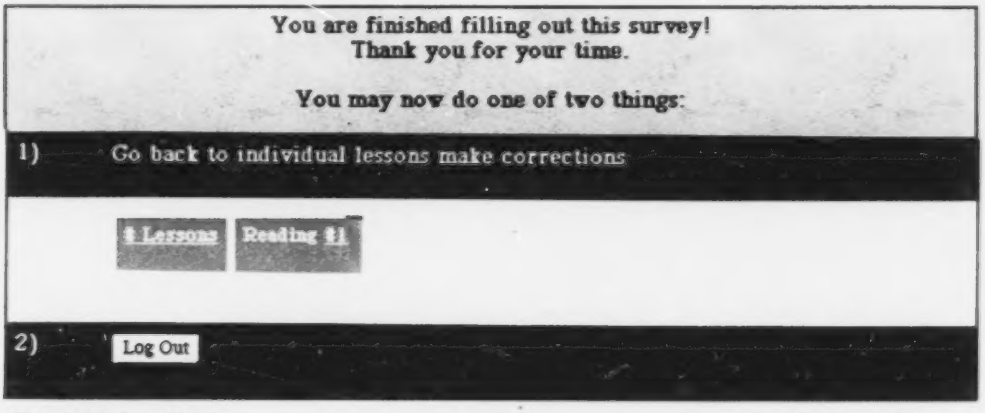

You are finished filling out this survey!
Thank you for your time.
You may now do one of two things:
1)
Go back to individual lessons make corrections
Reading #1
2)
Log Out

Figure 29: Log Out

BEST COPY AVAILABLE

CTP Working Papers

The Center's Working Paper Series allows researchers to release interim results of their ongoing investigations when these are sufficiently well developed to be informative to the field. Also, it gives education and policy audiences the opportunity to make use of these "works in progress" in a timely manner. We warmly encourage feedback on these documents to help us refine them further in preparation for final reports of our research. Papers in this series, as well as CTP Research Reports and Policy Briefs, are available for download from the Center's website: **www.ctpweb.org**

Center Team

Principal Investigators and Co-Principal Investigators

UNIVERSITY OF WASHINGTON
Michael Knapp, Center Director
James Banks
Pamela Grossman
Margaret Plecki
Sheila Valencia

STANFORD UNIVERSITY
Linda Darling-Hammond
Milbrey McLaughlin
Joan Talbert

UNIVERSITY OF MICHIGAN
David Cohen
Deborah Ball

UNIVERSITY OF PENNSYLVANIA
Thomas Corcoran

Researchers at Other Institutions
Barnett Berry, University of North Carolina
Richard Ingersoll, University of Georgia
David Monk, Pennsylvania State University
Edward Silver, University of Pittsburgh
Jon Snyder, University of California at Santa Barbara
Neil Theobald, Indiana University

Contact Information

Michael S. Knapp, Center Director
Miller Hall M201, College of Education
University of Washington, Box 353600
Seattle, WA 98195-3600
Email: mknapp@u.washington.edu

Michele C. Ferguson, Center Manager
Miller Hall 203C, College of Education
University of Washington, Box 353600
Seattle, WA 98195-3600
Phone: (206) 221-4114
FAX: (206) 616-6762
Email: ctpmail@u.washington.edu

Web Address
http://www.ctpweb.org